

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 09:27:12 ; Search time 1854.65 Seconds

(without alignments)  
16507.481 Million cell updates/sec

Title: US-09-805-311-1

Sequence: 1 cacgagaatagctcgcgcc.....aaaaaaaaaaaaaaaa 1463

### Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

post-processing:	Minimum Match	0%
	Maximum Match	100%

Database

```

1:  gb.ba:*
2:  gb.htg:*
3:  gb.in:*
4:  gb.om:*
5:  gb.ov:*
6:  gb.pat:*
7:  gb.ph:*
8:  gb.pl:*
9:  gb.pr:*
10: gb.ro:*
11: gb.sis:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sis:*
28: em.un:*
29: em.vi:*
30: em.htg_hum:*
31: em.htg_in:*
32: em.htg_other:*
33: em.htgo_in:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Query					Description
No.	Score	Match	length	DB	ID	
-----						

1	1463	100.0	1463	6	AR152403	Sequence
2	1421.8	97.2	1468	6	AR152406	Sequence
3	1409.8	96.4	1541	6	AR152404	Sequence
4	1350.2	92.3	1381	6	AR152405	Sequence
5	927.4	63.4	1354	8	AB021666	Sequence
6	328.6	22.3	1469	5	XL08141	Sequence
7	308.6	21.1	1461	5	XL064563	Sequence
8	307	21.0	1357	5	AF063397	Sequence
9	305.4	20.9	1149	5	AF063637	Sequence
10	305.4	20.9	2163	10	BC010203	Mus muscu
11	305.4	20.9	166088	2	AC026762	Mus muscu
12	302.8	20.7	1757	2	HSRAD2	Sequence
13	302.8	20.7	2031	9	AC000323	Sequence
14	302.8	20.7	138627	2	AP000551	Sequence
15	302.8	20.7	168035	2	AP002360	Sequence
16	302.8	20.7	185035	3	AC004770	Sequence
17	302.8	20.7	196089	8	AC004228	Sequence
18	302.6	20.7	1144	5	HUMEN4	Sequence
19	299	20.4	1146	5	HS058602	Sequence
20	292.4	20.0	39563	9	BS073629	Sequence
21	288.2	19.7	2192	10	AF281018	Rattus no
22	279.8	19.1	2033	10	MUSFEN1X	Sequence
23	278.2	19.0	6361	10	AY014962	Sequence
24	274.6	18.8	2114	8	SC1RL13C	Sequence
25	274.6	18.8	10660	6	SPAC36	Sequence
26	259.4	17.7	2299	8	SPRAD2	Sequence
27	259.4	17.7	28491	6	SPAC36	Sequence
28	259	17.7	151865	2	AL0356236	Sequence
29	259	17.7	179461	2	AC019177	Sequence
30	257.4	17.6	173256	2	AC073301	Sequence
31	257.4	17.6	178806	2	AC013557	Sequence
32	252.6	17.3	37688	2	AL513222	Sequence
33	252.6	17.3	95850	2	DMG5G003	Sequence
34	252.6	17.3	95850	2	AC020063	Sequence
35	252.6	17.3	134985	3	AC004641	Sequence
36	252.6	17.3	160338	3	AC099021	Sequence
37	252.6	17.3	270306	3	AE003805	Sequence
38	241.8	16.5	194322	3	AC024791	Sequence
39	241.8	16.5	319857	2	AC006782	Sequence
40	237	16.2	2783	3	AF093702	Sequence
41	237	16.2	234112	3	PFMAL42	Sequence
42	236.6	16.2	1953	3	AF278764	Sequence
43	236.6	16.2	194439	3	CNS07BEA	Sequence
44	236.6	16.0	120848	3	AC007801	Sequence
45	159.6	10.9	1023	6	AR062093	Sequence

## ALIGNMENTS

RESULT 1		
AR152403	1463 bp	DNA
LOCUS		linear
DEFINITION	Sequence 1 from patent US 6235257.	PAT 08-AUG-2001

Query Match	100.0%;	Score 1463;	DB 6;	Length 1463;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1463;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

OY 1 CACGAGATAGCTCGCGCGCGGGTTCTTGCGCCACTCCGGCTCAGCCCGCGCCGCCA 60
    |||
    |||
    |||
Db 1 CAGGAAATAGCTCGCGCGCGGGTTCTTGCGCCACTCCGGCTCAGCCCGCGCCGCCA 60
OY 61 CCGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
    |||
    |||
    |||
Db 61 CCGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
OY 121 AATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
    |||
    |||
    |||
Db 121 AATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
OY 181 GTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
    |||
    |||
    |||
Db 181 GTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
OY 241 ACTGTCACAAATGAGCTGTGGAAGTCAGTCTGATTTGCAAGGAATGTTCAACCGGACA 300
    |||
    |||
    |||
Db 241 ACTGTCACAAATGAGCTGTGGAAGTCAGTCTGATTTGCAAGGAATGTTCAACCGGACA 300
OY 301 ATAAGATTACTGGAAGCGGGGAATCAAGCCAGTTTATGTTTATGATGCAAGCCTCTGAT 360
    |||
    |||
    |||
Db 301 ATAAGATTACTGGAAGCGGGGAATCAAGCCAGTTTATGTTTATGATGCAAGCCTCTGAT 360
OY 361 ATGAAGAAACAAGCTGTGTTGTTAAAGATCTCAAAAAAGATGATGCAACCAAGATCTG 420
    |||
    |||
    |||
Db 361 ATGAAGAAACAAGCTGTGTTGTTAAAGATCTCAAAAAAGATGATGCAACCAAGATCTG 420
OY 421 ACTGAGCAGTAGAGTAGAGATGAAGATGCGATTGAAAAATTTAGCAAGAGACTGTA 480
    |||
    |||
    |||
Db 421 ACTGAGCAGTAGAGTAGAGATGAAGATGCGATTGAAAAATTTAGCAAGAGACTGTA 480
OY 481 AAGGTCACAGGCAACACACAGGAAGATTGTAACGCGCTATTAGACTTATGGGGTTCCT 540
    |||
    |||
    |||
Db 481 AAGGTCACAGGCAACACACAGGAAGATTGTAACGCGCTATTAGACTTATGGGGTTCCT 540
OY 541 GTTGTAGAGGAGACCTTGCGAAGCAGGAAGCAGATGTGACGCGCTTGTCATTAAGATAG 600
    |||
    |||
    |||
Db 541 GTTGTAGAGGAGACCTTGCGAAGCAGGAAGCAGATGTGACGCGCTTGTCATTAAGATAG 600
OY 601 GTGTCGCTGTGCTTCTGGAAGATATGACCTCCTTACTTTTGGGGCTCCAGCGTTCCTT 660
    |||
    |||
    |||
Db 601 GTGTCGCTGTGCTTCTGGAAGATATGACCTCCTTACTTTTGGGGCTCCAGCGTTCCTT 660
OY 661 CGTCTTTTAATGATGCCAAGTCCAGGAATACCTGTGATGGAATTTGATGTTGCCAAG 720
    |||
    |||
    |||
Db 661 CGTCTTTTAATGATGCCAAGTCCAGGAATACCTGTGATGGAATTTGATGTTGCCAAG 720
OY 721 GTTTTGGAGAGCTTGAAGTCCAGTCCAGTTCATTTGATGATCCTGTGTGA 780
    |||
    |||
    |||
Db 721 GTTTTGGAGAGCTTGAAGTCCAGTCCAGTTCATTTGATGATCCTGTGTGA 780
OY 781 TGTGACTATTGTGATAGCATCAAGGTATCGGGGGGCAACAGCTCTCAAACTTATTCCT 840
    |||
    |||
    |||
Db 781 TGTGACTATTGTGATAGCATCAAGGTATCGGGGGGCAACAGCTCTCAAACTTATTCCT 840
OY 841 CAACTATGGTTCATGAAGAATCTTTGGAAGATCTTAATAAAGACAGATATCAAAATTCCT 900
    |||
    |||
    |||
Db 841 CAACTATGGTTCATGAAGAATCTTTGGAAGATCTTAATAAAGACAGATATCAAAATTCCT 900
OY 901 GAGGACTGGCTTACCAAGAAAGCTGAGCGTTGTTCAAGAGACCTTAATGTCAATTTGAT 960
    |||
    |||
    |||
Db 901 GAGGACTGGCTTACCAAGAAAGCTGAGCGTTGTTCAAGAGACCTTAATGTCAATTTGAT 960
OY 961 ATTCTGAGCTAAATGAGCTGACCTGATGAGAGGGCTCTATAGTTCTCTGTTGTTAAA 1020
    |||
    |||
    |||
Db 961 ATTCTGAGCTAAATGAGCTGACCTGATGAGAGGGCTCTATAGTTCTCTGTTGTTAAA 1020
OY 1021 GATTAATGTTTCAAGAGATGCGGGTGACAAAGGCCATAGAGAATCAATCTGCCAAG 1080
    |||
    |||
    |||
Db 1021 GATTAATGTTTCAAGAGATGCGGGTGACAAAGGCCATAGAGAATCAATCTGCCAAG 1080

```

```

OY 1081 AATAATGCTGCGAAGAGACTGAGTCCCTTTTTCGAAGCCAACTGCCACACATGACGA 1140
    |||
    |||
    |||
Db 1081 AATAATGCTGCGAAGAGACTGAGTCCCTTTTTCGAAGCCAACTGCCACACATGACGA 1140
OY 1141 CCGCTAAACGGGAAGAGACTTCGATTAACCAAGCAGGAGCTGGGAGCAAGAAAACA 1200
    |||
    |||
    |||
Db 1141 CCGCTAAACGGGAAGAGACTTCGATTAACCAAGCAGGAGCTGGGAGCAAGAAAACA 1200
OY 1201 AAGCTGTGGAAGAAAGAAATTAATCTTGATGCTTATACACTACGACTACGAAAG 1260
    |||
    |||
    |||
Db 1201 AAGCTGTGGAAGAAAGAAATTAATCTTGATGCTTATACACTACGACTACGAAAG 1260
OY 1261 CAGCGGTGCGGTGATCACTTCCTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTT 1320
    |||
    |||
    |||
Db 1261 CAGCGGTGCGGTGATCACTTCCTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTT 1320
OY 1321 GGTAAATTTGCTCATGTTTCAAGCTGGGGTAGATTAGTTGTTGTTGAAGAGATTGGTG 1380
    |||
    |||
    |||
Db 1321 GGTAAATTTGCTCATGTTTCAAGCTGGGGTAGATTAGTTGTTGTTGAAGAGATTGGTG 1380
OY 1381 TACCAAGTAAACAAACTATGCGTGTGTTTAACTTCTTCTTGAAGTAAAAAAA 1440
    |||
    |||
    |||
Db 1381 TACCAAGTAAACAAACTATGCGTGTGTTTAACTTCTTCTTGAAGTAAAAAAA 1440
OY 1441 AAAAAAAAAAAAAAAAAAAAAA 1463
    |||
    |||
    |||
Db 1441 AAAAAAAAAAAAAAAAAAAAAA 1463

```

RESULT 2  
 LOCUS ARI52406 1478 bp DNA linear PAT 08-AUG-2001  
 DEFINITION Sequence 7 from patent US 6232527.  
 ACCESSION ARI52406  
 VERSION ARI52406.1 GI:15118456  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1478)  
 AUTHORS Mahajan, P. B.  
 TITLE Maize Rad2/FEN-1 orthologues and uses thereof  
 JOURNAL Patent: US 6232527-A 7 15-MAY-2001;  
 FEATURES  
 source 1..1478  
 /organism="unknown"  
 BASE COUNT 463 a 302 c 365 g 348 t  
 ORIGIN

Query Match 97.2%; Score 1421.8; DB 6; Length 1478;  
 Best Local Similarity 98.5%; Pred. No. 0;  
 Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

OY 7 AATAGCTCGCGCGCGGGTTCTTGCGCCACTCCGGCTCAGCCCGCGCGCGCGCGCG 66
    |||
    |||
    |||
Db 19 AATAGCTCGCGCGGGTCTTGCGCCACTCCGGCTCAGCCCGCGCGCGCGCGCGCG 78
OY 67 ACAAGCGCGCGACAGAGATGGGCAATCAAGGGTTTACCAAACTGCTGGCGGACATGGG 126
    |||
    |||
    |||
Db 79 ACAAGCGCGCGACAGAGATGGGCAATCAAGGGTTTACCAAACTGCTGGCGGACATGGG 138
OY 127 CCAAGCGCGATGAAGAGCAAGATTGAGAGCTACTGGCGCGCAAAATCGCGCTGCAG 186
    |||
    |||
    |||
Db 139 CCAAGCGCGATGAAGAGCAAGATTGAGAGCTACTGGCGCGCAAAATCGCGCTGCAG 198
OY 187 GCCAGCATGAGCATATACAGTTCTGATTTGATTTGGAAGAGCAGGCAATGGAATCTTC 246
    |||
    |||
    |||
Db 199 GCCAGCATGAGCATATACAGTTCTGATTTGATTTGGAAGAGCAGGCAATGGAATCTTC 258
OY 247 ACAATGAGCTGTGTAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACATTAAGA 306
    |||
    |||
    |||
Db 259 ACAATGAGCTGTGTAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACATTAAGA 318

```



QY	307	TTTGTGGAACGGGAGTTCAAAGCAGATTATGTTTTGATGTGCAAGCCCTCGATATTAAG	366
Db	319	TTACTGAGACGGGAAATCAACCCAGTTTATGTGTTTTGATGGCAAGCCCTCTGATATTAAG	378
QY	367	AAACAAGAGCTTCTAAAAAGTACTCAAAAAGAGATGATGCAACCAAGATCTGACTAG	426
Db	379	AAACAAGACTTTCCTAAAAGATACTCAAAAAGAGATGATGCAACCAAGATCTGACTAG	438
QY	427	GCAGTAGAGGTAGACATTAAGATGGGAGTTGAAAAATTGACCAAGAGAGACTGTAAAGCTC	486
Db	439	GCACTAGAGGTAGACATTAAGATGGGAGTTGAAAAATTGACCAAGAGAGACTGTAAAGCTC	498
QY	487	ACAAGGCAACACAAACAAGATTGTAAAGGCACTTTAAGACTTATGGGGGTTCCGTGTGA	546
Db	499	ACAAGGCAACACAAACAAGATTGTAAAGCACTATTAAAGCTTATGGGGGTTCCGTGTGA	558
QY	547	GAGGACACCTTCTAAGCAGAGAGAAATGGAGCCCTTGGATTAAGAGTAAGCTGTTC	606
Db	559	GAGGACACCTTCTAAGCAGAGAGAGATGTGAGCCCTTTGCAATTAAGATTAAGCTGTTC	618
QY	607	GCTGTGCTCTCAGAAGATATGACCTCCCTTACTTTTGGGGCTCCACGGTTCCTTGTCAT	666
Db	619	GCTGTGCTCTCAGAAGATTAAGAGCTCCCTTACTTTTGGGGCTCCACGGTTCCTTGTCAT	678
QY	667	TTAATGATCCAGATTTCACAAGAAATACCTGTGATGGAATTTGATGTGTCACAGSTTTTG	726
Db	679	TTAATGATCCAGATTTCACAAGAAATACCTGTGATGGAATTTGATGTGTTGCCAAGSTTTTG	738
QY	727	GAGGAGCTTGAACACTCAACATGACACAGTTTCATTGATTTTGTGCATCTGTGTGATGTGAC	786
Db	739	GAGGAGCTTGAACACTCAACATGACACAGTTTCATTGATTTTGTGCATCTGTGTGATGTGAC	798
QY	787	TATGTGATAGCATCAAAAGATATCGGGGGGGCAACAGCTCTCAACTATTTCGTCACAT	846
Db	799	TATGTGATAGCATCAAAAGATATCGGGGGGGCAACAGCTCTCAACTATTTCGTCACAT	858
QY	847	GGGTCCATACAAAGCATCTTGGAGAATCTTTATTAAGACACATATCAAAATTCCTGAGAC	906
Db	859	GGGTCCATACAAAGCATCTTGGAGAATCTTTATTAAGACACATATCAAAATTCCTGAGAC	918
QY	907	TGAGCTTACCAAGAAAGCTGCAGCTGTGTTCAGAGAGCCATAATGTACATTGGATATTCCT	966
Db	919	TGAGCTTACCAAGAAAGCTGCAGCTGTGTTCAGAGAGCCATAATGTACATTGGATATTCCT	978
QY	967	GAGCTAAATGAGACTGCACCTGATGAGAGGGGTCTCATTAAGTTTCTGTGTAAAAGATTAAT	1026
Db	979	GAGCTAAATGAGACTGCACCTGATGAGAGGGGTCTCATTAAGTTTCTGTGTAAAAGATTAAT	1038
QY	1027	GGTTTCAACAACAATGGGGTGACAAAAGGCCATAGAGAAGATCAAACTCGCCACAATAATAA	1086
Db	1039	GGTTTCAATGAAGTGGGGTGACAAAAGGCCATAGAGAAGATCAAACTCGCCACAATAATAA	1098
QY	1087	TGCTGCGAAGAGAGACTCGAGGCTCTTTTTCAGGCCACACTGCCACACATCAAGCCGCTA	1146
Db	1099	TGCTGCGAAGAGAGACTCGAGGCTCTTTTTCAGGCCACACTGCCACACATCAAGCCGCTA	1158
QY	1147	AAACGGAAGAGACTTCGATTAATAAACCAAGCAGCAGCTGCGCAACCAAAAAACAAGGCT	1206
Db	1159	AAACGGAAGAGACTTCGATTAATAAACCAAGCAGCAGCTGCGCAACCAAAAAACAAGGCT	1218
QY	1207	GGTGGAAGAAGAAATTAATCTTGGATGCTTGATGTACAACTCGAGCTCGAAGAACAGCAGG	1266
Db	1219	GGTGGAAGAAGAAATTAATCTTGGATGCTTGATGTACAACTCGAGCTCGAAGAACAGCAGG	1278
QY	1267	TGGCGTGATCACTTGGCTTATGATTTATTTAACTCCCTGTTTTAACTCAAGAGCTTGGTAAA	1326
Db	1279	TGGCGATGATCACTTGGCGCTGATTTATTTAACTCCCTGTTTTAACTCAAGAGCTTGGTAAA	1338
QY	1327	AGTTTGTCATGTTTCAAGCTGGGGGTAAGTTGATGTTGTTGAAGAGATGGGTGATCCAA	1386
Db	1339	AGTTTGCCCATGTTTCAAGCTGGGGGTAAGTTGATGTTGTTGAAGAGATGGGTGATCCAA	1398
QY	1387	GTAACAACAACCTATTCGCTGTTTTTTTACTCTTGCTGCTTGAAGTAAAAAATAAAAAA	1446

[illegible]

|||||  
Db 613 GAAATATGAGACTCCCTACTTTTGGGGCTCCAGGCTCCTTGCTATTTAATGATCCA 672  
Qy 679 AGTTCCAGAAATACCTGTGATGGAATTTGATGTTGCCAAGCTTTGGAGAGCTTGA 738  
Db 673 AGTTCCAGAAATACCTGTGATGGAATTTGATGTTGCCAAGCTTTGGAGAGCTTGA 732  
Qy 739 CTCACATGAGACAGTTCATGATTTGTCATCTGTGATGATGATGATGATGATGATG 798  
Db 733 CTCACATGAGACAGTTCATGATTTGTCATCTGTGATGATGATGATGATGATGATG 792  
Qy 799 ATCAAAAGTATCGGGGGGCAAAAGCTCTGAATTTATTCGTACATGAGCTTGA 858  
Db 793 ATCAAAAGTATCGGGGGGCAAAAGCTCTGAATTTATTCGTACATGAGCTTGA 852  
Qy 859 AGCATCTTGAGATCTTAATAAAGACAGATATCAATTCCTGAGAGCTGGCTTACCA 918  
Db 853 AGCATCTTGAGATCTTAATAAAGACAGATATCAATTCCTGAGAGCTGGCTTACCA 912  
Qy 919 GAACTGAGAGCTTGTTCAGAGAGCTTAATGTCACATTTGATTTCTGAGCTAAATG 978  
Db 913 GAACTGAGAGCTTGTTCAGAGAGCTTAATGTCACATTTGATTTCTGAGCTAAATG 972  
Qy 979 ACTGACCTGATGAGAGAGCTTCATAGTTTCTGTTAAAGATGATGTTTCAACGAA 1038  
Db 973 ACTGACCTGATGAGAGAGCTTCATAGTTTCTGTTAAAGATGATGTTTCAACGAA 1032  
Qy 1039 GATCGGCTGACAAAGGCTATAGAGAGATCAATTCGCCAAGATTAATCGTCGACGA 1098  
Db 1033 GATCGGCTGACAAAGGCTATAGAGAGATCAATTCGCCAAGATTAATCGTCGACGA 1092  
Qy 1099 AGACTCGAGTCTTTTTCAGAGCCACAGCCGACACATCGACATCGCTTAAAGAGAGAG 1158  
Db 1093 AGACTCGAGTCTTTTTCAGAGCCACAGCCGACACATCGACATCGCTTAAAGAGAGAG 1152  
Qy 1159 ACTTGGATTAACCAAGCAAGGAGCTGCGAGACAGAAACAAAGGCTGGTGAAGAAG 1218  
Db 1153 ACTTGGATTAACCAAGCAAGGAGCTGCGAGACAGAAACAAAGGCTGGTGAAGAAG 1212  
Qy 1219 AATTAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1278  
Db 1213 AATTAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1272  
Qy 1279 TTGCTTATGATTTATTAACCTGCTTTTAACTAGAGCTTTGGTAAAGTTGCTCATG 1338  
Db 1273 TTGCTTATGATTTATTAACCTGCTTTTAACTAGAGCTTTGGTAAAGTTGCTCATG 1332  
Qy 1339 TTTCAGGTGGGGTAACTAGTTGTTGTTGAAGAGATGGTGTACCAAGTAAACAACT 1398  
Db 1333 TTTCAGGTGGGGTAACTAGTTGTTGTTGAAGAGATGGTGTACCAAGTAAACAACT 1392  
Qy 1399 ATGCTGCTTTTAACTTCTTGTCTTGAAGA 1431  
Db 1393 ATGCTGCTTTTAACTTCTTGTCTTGAAGA 1425

RESULT 4  
ARI52405 1381 bp DNA linear PAT 08-AUG-2001  
LOCUS ARI52405 Sequence 5 from patent US 6232527.  
VERSION ARI52405  
KEYWORDS ARI52405.1 GI:15118455  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1381)  
AUTHORS Mahajan,P.B.  
TITLE Maize Rad2/FEN-1 orthologues and uses thereof  
JOURNAL Patent: US 6232527-A 5 15-MAY-2001;  
FEATURES Location/Qualifiers  
1..1381  
/organism="unknown"

BASE COUNT 441 a 269 c 346 g 325 t  
ORIGIN  
Query Match 92.3%; Score 1350.2; DB 6; Length 1381;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1355; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Qy 49 CGCGGCCCGCCACCGCCGACAGCCGCGGAGAGATGGGATCAAGGTTTGAGAAA 108  
Db 1 CGAGCCACGCGGTCCGCGCCAGCCGCGGAGAGATGGGATCAAGGTTTGAGAAA 60  
Qy 109 CTGCTGCGGACAAATGCGCCCAAGGCGATGAAGAGACAGAGCTGAGAGCTTGGCC 168  
Db 61 CTGCTGCGGACAAATGCGCCCAAGGCGATGAAGAGACAGAGCTTGGAGCTTGGCC 120  
Qy 169 CGCAAAATTCGCGCTGAGAGCGCAGCATGAGCATATACCATTCCTGATGATGAGAG 228  
Db 121 CGCAAAATTCGCGCTGAGAGCGCAGCATGAGCATATACCATTCCTGATGATGAGAG 180  
Qy 229 ACAGCATGGAACCTCTACAAATGAGAGCTGGTGAAGTCACTAGTATTCGACGAATG 288  
Db 181 ACAGCATGGAACCTCTACAAATGAGAGCTGGTGAAGTCACTAGTATTCGACGAATG 240  
Qy 289 TTCAACCGGACATTAAGATTAAGTGAAGGCGGAGATCAAGCCAGTTTATGTTTGATGGC 348  
Db 241 TTCAACCGGACATTAAGATTAAGTGAAGGCGGAGATCAAGCCAGTTTATGTTTGATGGC 300  
Qy 349 AAGCTTCTGATATGAAGAAACAGAGCTTGTCTAAAGATCTCAAAAAGAGATGATCA 408  
Db 301 AAGCTTCTGATATGAAGAAACAGAGCTTGTCTAAAGATCTCAAAAAGAGATGATCA 360  
Qy 409 ACCAAATGTCAGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468  
Db 361 ACCAAATGTCAGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
Qy 469 AAGAGACTTAAAGGTCAACAGGCAACAGCAAGAGATTTGAAGGCTATTAAGACTT 528  
Db 421 AAGAGACTTAAAGGTCAACAGGCAACAGCAAGAGATTTGAAGGCTATTAAGACTT 480  
Qy 529 ATGGGGTTCCTGTTGAGAGGCACTTCTGAAAGCAAGAGAGAGAGAGAGAGAGAGAG 588  
Db 481 ATGGGGTTCCTGTTGAGAGGCACTTCTGAAAGCAAGAGAGAGAGAGAGAGAGAGAG 540  
Qy 589 ATTAACGATTAAGGTGTTGCTGTTGCTTCAAGAGATGAGATGCTTCTTACTTTGGGCT 648  
Db 541 ATTAACGATTAAGGTGTTGCTGTTGCTTCAAGAGATGAGATGCTTCTTACTTTGGGCT 600  
Qy 649 CCAAGGTTCTTGTGATTTAATGATCCAGTTCACAAAGATTAACCTGATGATGATTT 708  
Db 601 CCAAGGTTCTTGTGATTTAATGATCCAGTTCACAAAGATTAACCTGATGATGATTT 660  
Qy 709 GATGTCGCCAAGCTTTTGGAGAGCTTGAACCTCACCATGAGCAGCTTATGATTTGTGC 768  
Db 661 GATGTCGCCAAGCTTTTGGAGAGCTTGAACCTCACCATGAGCAGCTTATGATTTGTGC 720  
Qy 769 ATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 828  
Db 721 ATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
Qy 829 AACTTATGTCATCAATGAGGTCATAGAGAACATCTTGAAGATCTTAATAAAGACGA 888  
Db 781 AACTTATGTCATCAATGAGGTCATAGAGAACATCTTGAAGATCTTAATAAAGACGA 840  
Qy 889 TATCAATTCCTGAGAGCTTGAAGAGAGCTCGACGCTTGTTCAGAGAGCTTAAT 948  
Db 841 TATCAATTCCTGAGAGCTTGAAGAGAGCTCGACGCTTGTTCAGAGAGCTTAAT 900  
Qy 949 GTCAATGATATTCCTAGATTAATGATGATGATGATGATGATGATGATGATGATGATG 1008  
Db 901 GTCAATGATATTCCTAGATTAATGATGATGATGATGATGATGATGATGATGATGATG 960  
Qy 1009 TTCTGTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1068  
|||||

Db 961 TTCTGGTAAAGATATGTTTCAACGAGAGTGGGTGACAAAGCCATAGAGAAATC 1020  
Qy 1069 AATCTGCCAAGATTAATCTGCCAAGAGACTCCAGTCTCTTTTCAAGCCACTGCC 1128  
Db 1021 AATCTGCCAAGATTAATCTGCCAAGAGACTCCAGTCTCTTTTCAAGCCACTGCC 1080  
Qy 1129 ACCACATGACGACCGCTTAAACGAGAGACTTGCATTAACAGCAAGGACGCTGG 1188  
Db 1081 ACCACATGACGACCGCTTAAACGAGAGACTTGCATTAACAGCAAGGACGCTGG 1140  
Qy 1189 AACAGAAACAAAGGCTGTGGAAGAGAAATATCTGATGCTGTGATGACAACTA 1248  
Db 1141 AACAGAAACAAAGGCTGTGGAAGAGAAATATCTGATGCTGTGATGACAACTA 1200  
Qy 1249 GCACATCAAGAGACGCGTGGCTGATCATTGCTTATTTTACCTCTCTTTT 1308  
Db 1201 GCACATCAAGAGACGCGTGGCTGATCATTGCTTATTTTACCTCTCTTTT 1260  
Qy 1309 ACTGACGCTTGTGTAAGTGTGCTCATGTTTCAAGCTGGGTAAGTGTGCTTTG 1368  
Db 1261 ACTGACGCTTGTGTAAGTGTGCTCATGTTTCAAGCTGGGTAAGTGTGCTTTG 1320  
Qy 1369 AAGAGATTGGTGTACCAAGTACAAACTTATGCTGCTTTT 1411  
Db 1321 AAGAGATTGGTGTACCAAGTACAAACTTATGCTGCTTTT 1363

RESULT 5  
AB021666 1354 bp mRNA linear PLN 25-MAR-2000

LOCUS AB021666  
DEFINITION Oryza sativa OsFEN-1 mRNA for FEN-1, complete cds.  
ACCESSION AB021666  
VERSION AB021666.2 GI:4587224  
KEYWORDS FEN-1; endonuclease.  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (sites)  
Kimura,S., Ueda,T., Hatanaka,M., Takenouchi,M., Hashimoto,J. and  
Sakaguchi,K.  
Plant homologue of flap endonuclease-1: molecular cloning,  
characterization, and evidence of expression in meristematic  
tissues  
Plant Mol. Biol. 42 (3), 415-427 (2000)  
2 (bases 1 to 1354)  
20256470  
Kimura,S., Hashimoto,J. and Sakaguchi,K.  
Direct Submission  
Submitted (21-DEC-1998) Seisuke Kimura, Science University of  
Tokyo, Dept. of Applied Biological Science, 2641 Yamazaki, Noda,  
Chiba 278-8510, Japan (E-mail:j3498703@ed.noda.sut.ac.jp,  
Tel:81-471-24-1501(ex.3419), Fax:81-471-23-9767)  
On Apr 17, 1999 this sequence version replaced gi:4062866.  
Location/Qualifiers  
1..1354  
/organism="Oryza sativa"  
/db\_xref="taxon:4530"  
67..1209  
/gene="OsFEN-1"  
67..1209  
/gene="OsFEN-1"  
/note="endonuclease"  
/codon\_start=1  
/product="FEN-1"  
/protein\_id="BA36171.1"  
/db\_xref="GI:4587225"  
/translation="MGIKGLTKLADNAPRAMEKQFESYFGRIAVDASMSIYOLLI  
VGRTEHETLNEAGVYSHLQGFNRTILFLAGIKPYVDEGKPPDLKQELAKRY  
SKRDATKELTEAVEEGDKDAIEKFSKRYKVTIKONNEBECKRLRLMGVVEAPCEA  
EACCAALCINDMVAVASSEMDSLTFEAPRLRHIDPSSKKIPIVEFVAVLELE  
LTMDPDLCLTSCDYCDISKIGGOTALKLRONGSIESILEINIKRQIPEDWP

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT  
FEATURES  
source

gene  
CDS

YOEARRLFKEPNVTLIDIPELKMNAPDEGLVEFLVKEGPNODRVTKAIEIKFAKNN  
SSQGRLESEFRRPVSTSVPLKRDISEKPTKAVANKTKTGAGGKKK"  
polyA\_site 1354  
/note="17 A nucleotides"  
BASE COUNT 415 a 286 c 331 g 322 t  
ORIGIN  
Query Match 63.4%; Score 927.4; DB 8; Length 1354;  
Best Local Similarly 86.3%; Pred. No. 1,2e-218;  
Matches 1037; Conservative 0; Mismatches 161; Indels 3; Gaps 1;  
55 CGGCCACCCGGCAGCGCGCGGAGAGATGGGCTCAAGGTTTGAGCAAACTGCTG 114  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 37 CGCCGCTCGCCGATCCGTGGAGGCGAGATGGGCTCAAGGTTTGAGCAAACTGCTG 96  
115 GCGCAAAATGCGCCCAAGGCGATGAGAGAGAGTTCGAGAGCTTCTGCGCCGCA 174  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 97 GCGGCAATGCGCCCAAGGCGATGAGAGAGAGTTCGAGAGCTTCTGCGCCGCGC 156  
175 ATGCGCGTGCAGCGCGAGCATGATATACAGTCTCTGATGATGAGAGAGCGC 234  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 157 ATGCGCGTGCAGCGCGAGCATGATATACAGTCTCTGATGATGAGAGAGCGC 216  
235 ATGGAACCTCTCAAAATGAAGAGTGTGAAGTCACTAGTCTTTCGAAGAGATGTTCAAC 294  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 217 ATGGAACCTCTCAAGAGAGAGCGCGTGAAGTCACTAGTCTTTCGAAGAGATGTTCAAC 276  
295 CGGCAATTAAGTATTACTGAGAGCGGGAATCAAGCAGTATGTTTGTGAGCAAGCCT 354  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 277 CGGCAATTAAGTATTACTGAGAGCGGGAATCAAGCAGTATGTTTGTGAGCAAGCCT 336  
355 CCGATATGAAGAAACAAGAGCTGCTTAAAGATATCAAAAAGAGATGATGCAACAAA 414  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 337 CCGATATGAAGAAACAAGAGCTGCTTAAAGATATCAAAAAGAGATGATGCAACAAA 396  
415 GATCTGACTGAGGAGTGAAGTGAAGATGAAGATGCAATGAAAAATTGAGCAAGAG 474  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 397 GAACTACAGAGGCGAGTGAAGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 456  
475 ACTGTAAGGCTCAAGAGCAACAGACAGATGTAAGCGCTTATTAAGACTTATGAGG 534  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 457 ACTGTAAGGCTCAAGAGCAACAGACAGATGTAAGCGCTTATTAAGACTTATGAGG 516  
535 GTTCTGTTGTGAGAGCACTTTCGAAGCAGACAGCAATGTGCACTTTCATTAAC 594  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 517 GTTCTGTTGTGAGAGCACTTTCGAAGCAGACAGCAATGTGCACTTTCATTAAC 576  
595 GATAAGTGTGCTGTTGCTTCAAGATATGAGCTCCCTTACTTTGGGCTCCACG 654  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 577 GATATGTTTATCGAGTTCATCAGATGAAGATATGAGCTTTCATTTGGAGCTCCACG 636  
655 TTCTTCTGCTATTTATGATCCAAAGTTCAGAAATACCTGTGATGATTTGATG 714  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 637 TTCTTCTGCTATTTATGATCCAAAGTTCAGAAATACCTGTGATGATTTGATG 696  
715 GCCAAGTGTGAGAGCTTGAACCTCAACATGAGCAGTTCATTTGTTGTCATCCG 774  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 697 GCGAAGTGTGAGAGCTTGAACCTCAACATGAGCAGTTCATTTGTTGTCATCCG 756  
775 TGTGATGTGACTATTTGATAGCATCAAGATATGCGGGGCGCAACAGCTGGAACCTT 834  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 757 TGTGATGTGACTATTTGATAGCATTAAGATATGAGAGTCAAAACGCTGGAACCTT 816  
835 ATTGCTCAACATGGGTCCATAGAAAGCATTTTGGAAAAATTAAGCAATATCA 894  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 817 ATTGCTCAACATGGGTCCATAGAAAGCATTTTGGAAAAATTAAGCAATATCA 876  
895 ATTGCTGAGAGCTGCTTACCAAGAGCTGAGCGTTTTCAGAGCCTTAATGTGACA 954  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
Db 877 ATTCTGAGGAGCTGCTTATCAAGAGCTGAGCGTTTTCAGAGCCTTAATGTGACA 936  
955 TTGGAATTTCTGAGCTTAAGATGAGCAGCTGATGAGAGGCTTCATTAATGCTC 1014  
11111 11111 11111 11111 11111 11111 11111 11111 11111 11111



Accession	Sequence	Length
D6	GTGAGCTGAGACCCGATGAGAGAGGAGGAGCTGTTGACCTTCATGTGTGGGAGAAAGCAGTTTCAG	1069
D6	1010 GTGAGCTGAGACCCGATGAGAGAGGAGGAGCTGTTGACCTTCATGTGTGGGAGAAAGCAGTTTCAG	1069
QY	1035 CGAAGATCGGGGTGACAAACAGCCCATGAGAAATGCAATATGCGCAAGATTAATCTGTGCA	1094
D6	1070 TGAAGACCGCATTCACCAATATGGGGCCCAAAAATGTTGGCCAAAGACCGTCACAGCGACACGCA	1129
QY	1095 AGGAAGACTCGAGTCTTTTTCACACCAACCGCCACCA	1132
D6	1130 GGGCCGGCTGGATGACTTTTTCAAAGTAGTACGTGATCTCA	1167

RESULT	7
Locus	XLU64563
Definition	XLU64563
Accession	Xenopus laevis 5' nucleosome RNA, complete cds.
Version	U64563.1 GI:1490869
Keywords	African clawed frog.
Source	
Source organism	
Source project	
Reference	VRT 26-JUL-1999

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
NUMBER

1 (bases 1 to 1461)  
Bibliova,M., Wu,B., Chi,E., Kim,K.H., Trautman,J.K. and Carroll,D.  
Characterization of FEN-1 from *Xenopus laevis*. cDNA cloning and  
role in DNA metabolism  
J. Biol. Chem. 273 (51), 34222-34229 (1998)  
99069415  
0957084

REFERENCE 2 (bases 1 to 1461)  
 AUTHORS Bibliova,M., Chi,E., Wu,B., Kim,K.-H. and Carroll,D.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-JUL-1996) Biochemistry, U. of Utah Med. Sch., 50 N.  
 Medical Dr., Salt Lake City, UT 84132, USA  
 FEATURES Location/Qualifiers

```

SOURCE
1. 1401
/organism="Xenopus laevis"
/db_xref="taxon:8355"
11. 1259
/definition="stimulated to produce Ovarian follicle growth"
CDS

```

```

during DNA replication and in base excision repair;
specifically removes single-stranded 5' sequence"
/note="similar to FEN-1, Dnaesiv, Mfl, Rad27, Rth1, and
rad2"
/codon_start=1
/product="5, nuclease xFEN1"
/protein_id="AAB06176.1"
/db_xref="gi:1490870"
/training="MGHGLAKLADYAPAIKEHDIKSYFGKRYAVDASMCITYELL
AVROGNTLQNEGGFTTSHLMGFRTIRAVEGIIKPYVFOCKPOMKSGELARSR
AAAEELKLEAAEEGEEVETNEKFTKRLVAVYQNHDECKLTLLMGITVYAPCEAA
ATCALVAKGAYVAAATFEDMDALFECPVLLRLTPASEKKKPIIDIEPHINRYIDIG
THEQVDCILIGSDYCEETITGIGSKRAIDILQKHTIDIEINDIDIKKYPPEWMLL
KKAHLPLEPEVLPDITELKMIPEDEEGVLAIVMGCEKQSEDRINGKAKTLAKNRQ
STQGRILDFEYVATGVSSTFKREKESGSAKKRAKGTGPAKCFKRG"

```

	Query Match	21.1%	Score 308.6;	DB 5;	Length 1461;
	Best Local Similarity	55.7%	Pred. No. 1.6e-65;		
	Matches 634;	Conservative	0;	Mismatches 499;	Indels 6; Gaps 2
OY	85 ATGGGCATCAAGGCTTTGACGAAACTGCTGCGCGCAATGCGCCCAAGGCCATGAGAG	144			
Db	111 ATGGGATTTACAGGTTTGGCCAAACTTATTCGAGATGTGGACCTGACGATCAAGAAG	170			
OY	145 CAGAGCTTCAGAGCTACTTGGCGCGCAAAATCGCGCTCGACGCGCAGCATGAGCATATAC	204			
Db	171 CATGATTTAAAAAGTTACTTTGGTCGAAAAGGTGGCGCATGTCTTCATGTCGATTTAT	230			
OY	205 CAGTCTCGATTAAGTACGTGGAAGGACAGGCATGGAATCTCCACAATGGAAGCGTGGA	264			

Db	231	CAGTTCTTATATGCAAGTGGAGACAGAGATGGCA--ACAGCGTCGCAAAACGAGAGAGTGA	287
OY	265	GTACACTAGTCATTTTGGCAAGCAATGTTCACCCGGGCATTAAGATTACTGGACGGCAATC	324
Db	288	ACCACAGGCCATCTTATGAGGATGTTTTTACCGTACCATTTGCGATGGCATGGCATGGCATG	347
OY	325	AAAGCGATTATGTTTTTGGATGGCAAGCGCTCCGTATATGAAGAAACAGAGCTTCTATAA	384
Db	348	AAACCGATGATGTGTTTATGGGAGCCACACAGATGAAGTCCGGTAACTAGCTATAA	407
OY	385	AGATACTCAAAAAGATGATGCAACCAAGAAACGTCAGTACAGGACAGTATAGATAGAGAT	444
Db	408	CGGATGGACCGTATGGGCGGAGGACAGAGAGTTGCTTGAACCTCCAGAGAGGCTGGAGAA	467
OY	445	AAAGATGCGATTGAAAATTTGAGCAAGAGAGACTGTAAAGTGCACAGGACACAAACGAA	504
Db	468	GTGGAGAAACATTTGAGAGAGTGCCTAAAGACGTTGTCAAGGTTACCAAGACGACATAGAG	527
OY	505	GATTGTAACCGGCTPTTAAAGATTGTGGGGTCTCTGTGTAGAGGACCTCTGTAGAGA	564
Db	528	GAGTGCAGAAATTAATTAACCTTAATGAGGACTTCTTATGTGATGACACCTCTGTGAAGT	587
OY	565	GAAAGCAAGATGTGCACAGCCCTTTGCTATAACGATTAAGGTGTCTGTCTGTGACAGAT	624
Db	588	GAGCCACAGCTGTGCTCCCTAGTAAACACAGAGAAACATGATGCTACAGCAACACAGAC	647
OY	625	ATGAGCTCCCTACTTTTGGGGGCTCCACAGTTCCTGCTGCTATTAATGATGCCAAGTTCC	684
Db	648	ATGAGATGCTCTGACCTTTTGGCACCCCTGTCTCTTTAGGCACTCCACAGCAAGCAAGT	707
OY	685	AAGAAATACCTGTGATGCAATTTGATGTTGCCAAGSTTTTGGAGAGAGCTTGCAATCCAC	744
Db	708	AAAAGCTGCTATTTCAGAGAGTCCATTTAAACCGTGTCAATTCAGGACATAGGTATACAA	767
OY	745	ATGAGACAGTCTATGATTTGTGCATCTGTGTGGATGTGATATTTGATTACATCAAA	804
Db	768	CATGACACATTTTGTGATCTCTGATCTCTGCTCGGAAGTACATCTGTGAACACATCCGT	827
OY	805	GGATCGGGGGGCAACAGCCTTGAACTTATTGCTAACATGGGTCATAGAAAGCATC	864
Db	828	GGCATTTGGACCCAAAGAGCCATTTACCTGATCGGTGAGATTAAGACCATTCAGCAGATC	887
OY	865	TTGAGACATCTTAATTAAGACGATATCAATTTCTGAGAGCTAGGCTTACCAAGAGCT	924
Db	888	ATTGATTAACATTCACCTCAAGAAATACCCCGTTCCGGAATTCGGCTGCACAGAGAGCA	947
OY	925	CGAGCCTTGTTCAAGAGCTTA--TGTCACTTTGGATTTCTCTAGCTTAAATGAGCT	981
Db	948	AAACATCTTTTCTTTGGAAACCAAGAACTACACTGACATCCAGGAGCGAAGTGGATC	1007
OY	982	GCACCTGATGAGAGAGGCTTCATTAACTTCTCGGTAAAGATTAAGTGTTCACAGCAAT	1041
Db	1008	GAGCCGAGTGAAGAAAGATTTGGTGGCTCTATGTTGTGGGAGAAACAGTTTAACTGAAGAT	1067
OY	1042	CGGGTGCACAAAGCCATAGAGAATCAATATCTGCGCAAGAATTAATCTGTGCAAGAGACA	1101
Db	1068	CGCATACGCAACGGGGCCAAAAGATTGGCCACAAGAACCGGCAAGGACAGCACACAGGGCCGG	1127
OY	1102	CTCGAGTCTTTTTCACAGCAACTCCACACATACAGACCGCTAAACCGGAAGGAGACT	1161
Db	1128	CTGGATGACTTTTTCAAAGTGAACCGGATCCGTTAGCTCACTAAAGAAAGAGCGCAGAA	1187
OY	1162	TTGGGATTAACAAAGCAGAGCAGCTCCGAAACAAGAAACAAAGCTGTGTGGAAGAAGAA	1220
Db	1188	TCCAAAGAGATCGGCTTAAGAAAGAAAGCAAAAGCTGAGAGTACCCCTGCGAGAAAGTTCAA	1246
RESULT 8			
AF065397			
LOCUS AF065397 1357 bp mRNA linear VFT 05-JAN-1999			
DEFINITION Xenopus laevis flap endonuclease 1 (FEN1) mRNA, complete cds.			
ACCESSION AF065397			
VERSION AF065397.1 GI:4106357			

[illegible]

D	b		516	GATTCGACGAATAATTCTTAACCCCTAATGGGCATTCCTTATGTGATCGACCCGTGAGACT	575
O	y		565	GAAGCAGAAATGTGCAAGCCCTTGATTAAGCATTAAGSTTTGCCCTTGTCCAGAAAGT	624
D	b		576	GAGGCCACAGTGTGCTCCCTTAGTAAGAACCAGGAAGATGATGCTGCAGCACACAGAAAGC	635
O	y		625	ATGGACTTCCTTACTTTTGTGGGCTCCACAGGTTCCTTCGTGCTATTAAATGATCCAACTTCC	684
D	b		636	ATGATGATCCCTTACCTTGTGGCCACCCCTTGTGCTCCTTGTAGGCACTCACAGACCGAAGCT	695
O	y		685	AAGAANAATCCGTGATGAGGAATTTGATGTTGCCAAGGTTTTGGAGGAGCTTGTAACACACC	744
D	b		696	AAAAGCTGCTTAATTCAGAGATTCATTTAAACCGTGTCTCATGAGACATAGGTATCACA	755
O	y		745	ATGGACCACTTCATTTGATTTGTCATCTGCTGTGATGATGACATTAATGTGATACATCAAA	804
D	b		756	CATGAACCAATTTGTGATCTCTGTATCTTCCTCGGAAGTACTACTATGAAACCATCTCGT	815
O	y		805	GGTATGGGGGGCAAACAGCTCTGAACCTTATTCGTCAACATGGGTCCATAGAAAGCATC	864
D	b		816	GGCATTGGACCCAAAGAGCCATTGACTGATCCGTTCAGATTAAGACCATTTGACGATGATC	875
O	y		865	TTCGAGATCTTAAATTAACACAGATATCAAAATTCCTTGAGACTGCTTACCAAGAGCT	924
D	b		876	ATTGATTAACATTTGATACCTCAAGAAATATACCCGCTCCGGAAAACTGGCTGCACAAGAGG	935
O	y		925	CGACGCTGTGTCAGAGAGCCTAA--TGTCAATATGGATATTCCTGTAGCTAAATAAGTACT	981
D	b		936	AAMCATCTTTCTTGGAACCAAGATATGTCGACTGACATCACCGAGCTGAAGTGATTC	995
O	y		982	GCACCTGATGAGAGGGTCTCATAGTTTCTTGTTAAAGATTAATGGTTTCAACGAGAAAT	1041
D	b		996	GAGCCGATGAAAGAGATTGGTCCCTTCAATGTGTGGGGGAAGACATTCATGTGAAGAC	1055
O	y		1042	CGGGTGCAAAAGSCCATAGAGAAGATCAAAATCTGCCACAAGATAAATATGTCGACAGSAGA	1101
D	b		1056	CGCATACGCAACGGGGCCAAAAGTTGGCCAAAGAACCGGCAAGGCAAGCAGCGAGCCCG	1115
O	y		1102	CTCGAGTCCCTTTTTCAGGCACATGCGCACACATCAGACCCGCTTAAACGGAAGAGACT	1161
D	b		1116	CTGATGACTTTTTCACAAATGACCGGATTCCTGTAGCTCACTAAATAAAGAGCGACAGA	1175
O	y		1162	TTCGATTAACAAGCAGAGCAGCTCGCAACAGAAACAAAGAGCTGTGTGAAAGAAAGA	1220
D	b		1176	TCCAAGAGATCCGCTTAAGAAAGAACCCAAAACTGGAGGTTACCTCTGCAGGAAAGTTCAA	1234
<hr/>					
RESULT 9					
AF036327					
LOCUS AF036327 1149 bp mRNA linear VRT 26-OCT-2006					
DEFINITION Xenopus laevis flap endonuclease 1 (FEN1) mRNA, complete cds.					
ACCESSION AF036327					
VERSION AF036327.1 GI:2674206					
KEYWORDS .					
SOURCE African clawed frog.					
ORGANISM Xenopus laevis					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;					
Xenopodinae; Xenopus.					
REFERENCE					
AUTHORS Kim K., Bide S. and Matsumoto Y.					
TITLE Involvement of flap endonuclease 1 in base excision DNA repair					
JOURNAL J. Biol. Chem. 273 (15), 8842-8848 (1998)					
MEDLINE 98204872					
PUBMED 9535864					
REFERENCE 2 (bases 1 to 1149).					
AUTHORS Kim K. and Matsumoto Y.					
TITLE Direct Submission					
JOURNAL Submitted (01-DEC-1997) Radiation Oncology, Fox Chase Cancer					
FEATURES Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA					
location/Qualifiers					
1..1149					
/organism="Xenopus laevis"					

/db\_xref="taxon:8355"  
 /cell\_type="ocyte"  
 1.1149  
 /gene="FEN1"  
 1.1149  
 /gene="FEN1"  
 /function="structure-specific nuclease"  
 /product="flap endonuclease 1"  
 /protein\_id="AAB8707.1"  
 /db\_xref="GI:2674207"  
 /translation="MGIHLAKLADVAPAIKHEHDKSPERKVAIDASMCIOFLI  
 AVRODNTIONEGETTSHLMFRTITMVEHGKIPVYEDSKPPQKSGSLARSE  
 RRAAEKLEAEAESEVENIEKFTRLVKVTKQNEBECKRLITLGLPYDAPCEAE  
 ATCAALVKAKVIAAATIEDMALFTGTPVLRHLAASEKKLPIDFHLNRYIDIGI  
 THROVDICILGSDYCEITRGISPKRAIDLIROKHTIDEIINDIDLKVPVBNMLH  
 KEAKHLEPEVVDITTELKWIPEDECELVAFMCEKOFSDRIENGAKKLAKNRQG  
 STQRLDFFKVTGVSSTKRKESESKAKKKAKTKGTGTPAKFRKRGK"

BASE COUNT 352 a 247 c 306 g 244 t  
 ORIGIN

Query Match 20.9%; Score 305.4; DB 5; Length 1149;  
 Best Local Similarity 55.5%; Pred. No. 9,4e-65;  
 Matches 632; Conservative 0; Mismatches 501; Indels 6; Gaps 2;

QY 85 ATGGCATCAAGGGGTTGACGAACCTGCTGGCGGACATCGGCCCAAGCGGATGAGGAG 144  
 DB 1 ATGGCAATTCACGGTTTGGCCAAACTTATCGCAGATGTGGACCTGAGCTATCAAGAAG 60  
 QY 145 CAGAAGTTCGAGAGCTACTTGGCCGCAAAATCGCTCGACAGCCCATGACATATATAC 204  
 DB 61 CATGATATATAAAGTACTTGTGTGCAAAAGGTGCGCATCGATCGTCTCATGCTATTTAT 120  
 QY 205 CAGTTCCTGATTGTGATGGAAGAGCATGGAACCTGCACAAATGAAAGCTGTGGA 264  
 DB 121 CAGTTTCTTATTTGGGTGAGACAGATGCA--ACACGCTGCAAAAGCAAGAGTGAA 177  
 QY 265 GTCACTAGTCTTTGCAAGATGTTCAACCGGACAAATGAGTACTGTGAAGCGGAGAT 324  
 DB 178 ACCACAACCATCTTATGGTATGTTTACGTCACATTCGCATGATGATGAGCAATGCGATC 237  
 QY 325 AAGCAGTTTATGTTTGTGATGAGCAAGCTCTCTATATGAAGAACAAGAGCTTCTAAA 364  
 DB 238 AAACCAAGTATGTTTGTGATGAGCAAGCGCCGACAGATGAAGTCCGTTGAACATGACAAA 297  
 QY 385 AGATACCTCAAAAGAGATGATGACCAACCAAGATGACTGAGCAGTAGAGTAGAGAT 444  
 DB 298 GCGACTGAGCGTAGGCGCGAGAGAGATGCTTGAAGCTGCCAGAGGCTCGAGAA 357  
 QY 445 AAAGATCGGATTGAAAATTTAGCAAGAGGACTGTAAAGTTCACAGAGCACAACAGAA 504  
 DB 358 GTGGGAACATTTGAAATTCATAAAGACTTCTCAAGGTACCAAGCAGCAACATGAG 417  
 QY 505 GATTTAAACGGCTATTAGACTTATGGGGGTCTCTGTTAGAGGACACTTCTGAGCA 564  
 DB 418 GAGTCAAGAAATTTACTTACCTTATGGGCAATTCCTTATGATGATGACCGGTGTGAAGCT 477  
 QY 565 GAAGCAAGATGTGACAGCCCTTTCATTAACATTAAGTGTTCGCTGTTCTTAAAGAT 624  
 DB 478 GAGGCGACGTGTGCTTGTAGTAAGGACAGGAAGTGAAGTGTGAGCAAGAAAGAC 537  
 QY 625 ATGGATCCCTTACTTTTGGGGCTCCAGGTTTCTGTCATTTAATGATGATCCAAAGTTTC 684  
 DB 538 ATGGATGCCCTTGACCTTTGGCACCCCTGTCTTATAGGACCTCACAGACAGCAAGACT 597  
 QY 685 AAGAAATACCTGTGATGGAATTTGATGTGCGCAAGTTTGGAGAGCTTGAAGTCAAC 744  
 DB 598 AAAAAGCGCTTATTCAGAGAGTTCATCTAAACGTCATTCAGGACATAGGATACACA 657  
 QY 745 ATGGAAGATTCATGATTTGTGATCCTGTGTGATGATGATTTGTGTATGATCAAA 804  
 DB 658 CATGAACATTTGTGATCTCTGTATCTTGTGCGAAGTAGCTACTGTGAACACCTTCGT 717

QY 805 GGTATCGGGGGCAACAGCTCTGAACCTTATTCGTCAACATGGGTGCATAGAAAGCATC 864  
 DB 718 GGCAATGGACCCCAAAAGAGCGATTTGACCTGATCCGTGAGATTAAGCACTTGGACGATC 777  
 QY 865 TTGGAGAACTTTAATAAAGACAGATATCAATTCCTGAGAGCTGGCTTACCAAGAGCT 924  
 DB 778 ATGTATAACATTTGACCTCAAGAAATATCCCGTCCGAAATCTGGTGGCAAGAGGCGCA 837  
 QY 925 CGAGCTTGTTCAGAGAGCCGTAA--TGTACATTTGATATTCCTGAGCTTAAATATGACT 981  
 DB 838 AAACATCTTTTCTTGAGAACCAAGAGTAGTGACATGACATCAACGACGCTGAAGTGCATC 897  
 QY 982 GCACCTGATGCGAGGCTTCATAGTTTCTGCTGTAAGATATATGTTTCAACAGACT 1041  
 DB 898 GAGCCCGATGAAGAGGATGTGGCCCTTATGATGTGGGAGAGAGAGATTTAGTGAAGT 957  
 QY 1042 CGGCTGCAAGGCCCTATAGAGAAATCAATTCGCCAAGATTAATCTGCGCAAGCAAA 1101  
 DB 958 CGCATACCCCAAGGGGCCCAAAAGTTGGCCCAAGAACCGGCAAGCAGCAGCAGGCGCG 1017  
 QY 1102 CTGAGTCTCTTTTCAAGCCACTGCCACCATCATGACACGCTTAAACGGAAGAGACT 1161  
 DB 1018 CTGATGACTTTTTCAAAGTGACCGGATCCGTTAGCTCAACTAAAGAAAGAGGAGAA 1077  
 QY 1162 TCGATTAACACAGCAGGCGAGCTGCGACACAGAAACAAAGGCTGGTGAAGAGAA 1220  
 DB 1078 TCCAAAGATCCGCTTAAAGAAAGCCAAACCTGAGAGTACCCTGACGAGAAAGTTCAA 1136

RESULT 10  
 BC010203 2163 bp mRNA linear ROD 22-OCT-2001  
 LOCUS Mus musculus, similar to flag structure-specific endonuclease,  
 DEFINITION BC010203.1 GI:16307327  
 ACCESSION BC010203  
 VERSION BC010203.1  
 KEYWORDS MGC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2163)  
 AUTHORS Strausberg,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-JUL-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 CONTACT: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Jeffrey Green M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Sequencing Center  
 Center code: BCM-HSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [villaloboscm.tmc.edu](mailto:villaloboscm.tmc.edu)  
 Villalobos, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia,  
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
 Muzny,D.M., Gibbs,R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 6 Row: d Column: 19.  
 Location/Qualifiers  
 1. 2163

FEATURES  
 source  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="MGC:7137 IMAGE:3158176"  
 /tissue\_type="mammary tumor, C3(1)-Tag model. Infiltrating  
 ductal carcinoma. 5 month old virgin mouse."



```

/clone_1lb="NCI_CGAP_Mam6"
/lab_host="DH108"
/notes="Vector: PCMV-S-SPORT6"
206..1348
CDS
/codon_start=1
/product="Similar to flag structure-specific endonuclease"
/protein_id="AAH10203.1"
/db_xref="GI:16307328"
/translation="MGHGLAKLADYAPSAIRENDIKSYFRKRAIDAMSITYOFLI
AVRGGVLDONAEGETEETSHLMGEYRIRRMENGIRPVYFDEKRPOLKSGELAKRSE
RRAREKOLDAOAGMEVEVEKFKRVLVYVTHKOHDECHLISLPIDAPSAE
ASCAIAKAKKAVYAAATREMDCLTFGSPVIMRLTASEAKLPLEIOPFHLSRYLOELG
NOEOPVLCILGSDICESIRGIDPKRAVYIDKHSIEITVIRLDSPYVPEMNLH
KEAOQLFEPEVLDPESEVLEIKWSPNEELVTKFCKQSEIRIKSVYKRLSKSHOG
STOGRLDFFKVTGSLSSAKRKREPEPKRPAKRAKTGGAGKFRFRGK"
BASE COUNT      540 a      528 c      619 g      476 t
ORIGIN
Query Match      20.9% Score 305.4; DB:10; Length 2163;
Best Local Similarity 56.3%; Pred. No. 1.1e-64;
Matches 642; Conservative 0; Mismatches 481; Indels 18; Gaps 3;
OY 85 ATGGCAGTCAAGGGTTTGCAGAACTGCTGGCGGACAAATGGCGCCAGCGCATGAAGAG 144
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 206 ATGGGAATTCAGGGCTTGCCAACTAATGTGATGTGGCCCCAGTGCATCGGTGAG 265
OY 145 CAGAACTTGAGAGCTACTTGGCGCCGCAAAATGCCGCTGACGCCAGCATGACATATAC 204
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 266 AATGACATCAAGAGCTACTTGTGTGCAAAAGTGGCCATGATGCCATGAGCATCTAC 325
OY 205 CAGTCCCTGATGTGTGGAGAGGAGCATGAAACTCTGACAAATGAAGCTGTGTAA 264
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 326 CAGTTCCTGATGTGTGTGTGTGAG--GTTGGGAGATGTGCTCAACAGAGAGAGGTGAG 382
OY 265 GTGACTAGTCAATTTGCAGAGAAATGTTCAACGACAAATAGATTACTGGAAGCGGAAATC 324
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 383 ACACAGAGCACTGATGTGGGATGTTCTACACACATCCGATGATGAGAAATGGCATC 442
OY 325 AAGCCAGTTTATGTTTGTATGAGCAAGCTCTGATATGAACAAACAGAGTTCTTAA 384
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 443 AACCCTGTGACTCTTTATGAGCAAAACACACAGCTAAATGAGGCGGCTGGCCAG 502
OY 385 AGATACTCAAAAAGAGATGATGCAACCAAGATCTGACTAGAGGAGTAGAGAT 444
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 503 CGCAGTGAAGAGGCGGCCGACGCTGAGAAAGCAACTCAGCAGGCTCAGGAGCTGGATG 562
OY 445 AAGATGCGATTGAAAAATGAGCAAGAGAGCTGTAAGGTCAAGCAAGCAACAA 504
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 563 GAGGAGAGAGGTGAGAAAGTTCAACAGAGGCTGTGCAAGGTCAACCAAGCAACATGAT 622
OY 505 GATTGTAACGGCTATTAGACTTATGAGGGGTCTGTTAGAGGAGCAGCTTCTGAGCA 564
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 623 GAGTGAACAACTGCTGAGAGCTGATGAGGATCTCTTACCTGATCAGCCAGCAGGAG 682
OY 565 GAAGCAAAATGTGACAGCCCTTTGCAATAAAGATAGGTGCTGCTGTTCTCAGAAAT 624
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 683 GAGGCGAGCTGTGCTGGCCGCGCAAAAGGCTGCAAAAGTATGCTGCGCCAGCAGAGAC 742
OY 625 ATGAGATCCCTTACTTTTGGGGCTCCAGAGTCTCTTCATTTATGATGATCCAGTCC 684
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 743 ATGAGATGCTCTCACTTTGAGCAGCCCGGTCTATGCGACACTTAAGTCTGAGTGGGCC 802
OY 685 AAGAAATACCTGATGATTTGATTTGATTTGATTTGAGGAGCTTGAAGTCAAC 744
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 803 AAGAGTGGCCATCAAGAGTTCATCTGAGCCGCTCTCTGAGAGAGTGGGTCTGAGAC 862
OY 745 ATGAGCAGCTCATTTGATTTGTGATCCTGTGTGATGATGATTTGTATGATCAACAA 804
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 863 CAGAGAGAGTTTGTGATCTGTGATCTGCTGGTGAAGCACTACTGCGAGACATCCGT 922
OY 805 GGTATCGGGGGGCAAGAGCTGGAACCTATTTCGTCACATGGGTGTCATGAAGACATC 864
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 923 GGCATTTGGGCCCAAGCGGCTGTGTGATCTCATCAACAAATTAAGCATCGAGAGATC 982

```

```

OY 865 TTGGAGATCTTAATAAAGACAGATATCAAAATTCCTGGAGAGCTGGCTTACCAAGACT 924
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 983 GTGAGGCGGCTGAGCCCGCAGAGTACCCGTTCCAGAACTGCTCTCCACAGGAACCC 1042
OY 925 CGACGCTGTTCAAGGAGGCTTAATG---TCACATTGATATTCTGAGCTAAATATGACT 981
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 1043 CAGCAGCTCTCTCTGAGCAGAGATCTGAGACCCAGAGTCTGTGAGACTGATGAGAGC 1102
OY 982 GCACCTGATGAGAGGCTCTCATFAGTTTCCTGTTAAAGATTAATGTTTCAACAGACAT 1041
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 1103 GAGCCAAATGAAGAAAGTGTGTCATTAATTTATGTGTGTGTAAGAACAGATTTCGAAAG 1162
OY 1042 CGGGTCAACAAGGCCATAGCAAGATCAAAATCTGCCAAGATAAATCTGTCGAAGAGAGA 1101
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 1163 CGAATTCGACAGTGGGTCAAGGCGCTGAGTAGAGCGCGCAGAGGAGCAGACCCAGGAGCGC 1222
OY 1102 CTCGAGTCTTTTTCAGGCCCACTGCCACCATCTACACACCCCTAAACGGAAGAGACT 1161
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 1223 CTCGATGATTTTCTCAAGGAGACAGGCTCCTCTCTCAGC-----TAAGCGC 1270
OY 1162 TCGGATTAACCAAGCAAGCAGCTGGACACAAAGAAAGAGCTGTGGAAGAGAGAA 1221
    ||||| || || || || || || || || || || || || || || || || || || ||
Db 1271 AAGGAGCAGAACCCAGAGGGGCTGCTAAGAGAGAAAGCAAGACTGGGGAGCCGGGAAG 1330
OY 1222 T 1222
Db 1331 T 1331

RESULT 11
AC026761      166088 bp      DNA      linear      HTG 25-JAN-2002
LOCUS      AC026761
DEFINITION      Mus musculus chromosome 19 clone RP22-325J22 strain 129/SVETAC1br,
WORKING DRAFT SEQUENCE, 25 unordered pieces.
ACCESSION      AC026761.10 GI:15217172
VERSION      AC026761.10 GI:15217172
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      house mouse.
ORGANISM      Mus musculus.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE      1 (bases 1 to 166088)
AUTHORS      Grills,G., Han,J., Montgomery,K.T., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Fereta,A.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
TITLE      Mouse High Throughput Sequencing
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 166088)
AUTHORS      Grills,G., Han,J., Montgomery,K.T., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Fereta,A.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
TITLE      Direct Submission
JOURNAL      Submitted (24-MAR-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
COMMENT      On Aug 21, 2001 this sequence version replaced gi:10937955.
-----Genome Center
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site:
http://wehanning.bwh.harvard.edu:9088/hpccg/jsp/hpccg/Sequence/mous
e.html
Contact: gnktm@capcod.bwh.harvard.edu
Center project name: ACD
-----Summary Statistics
Sequencing vector: pUC18; 108752
Chemistry: Dye-terminator Big Dye 100%
*Consensus quality: 156712 at least 920
*Consensus quality: 152694 at least 930
*Consensus quality: 147247 at least 940
Estimated insert size: agarose-FP - N/A
**Estimated insert size: 165608 - sum-of-contigs

```



Quality coverage: agarose-FP - N/A  
Quality coverage: 4 x in Q20 bases; sum-of-coverage estimation

NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 34556: contig of 34556 bp in length  
34557 34576: gap of unknown length  
34577 62727: contig of 28151 bp in length  
62728 62747: gap of unknown length  
62748 81216: contig of 18469 bp in length  
81217 81237: gap of unknown length  
81237 92870: contig of 11634 bp in length  
92871 92890: gap of unknown length  
92891 103627: contig of 10737 bp in length  
103628 103647: gap of unknown length  
103648 114697: contig of 11050 bp in length  
114698 114717: gap of unknown length  
114718 123385: contig of 8668 bp in length  
123386 123405: gap of unknown length  
123406 131815: contig of 8410 bp in length  
131816 131835: gap of unknown length  
131836 132037: contig of 202 bp in length  
132038 132057: gap of unknown length  
132058 138473: contig of 6416 bp in length  
138474 138493: gap of unknown length  
138494 144023: contig of 5530 bp in length  
144024 144043: gap of unknown length  
144044 144924: contig of 881 bp in length  
144925 144944: gap of unknown length  
144945 145191: contig of 247 bp in length  
145192 145211: gap of unknown length  
145212 148414: contig of 3203 bp in length  
148415 148434: gap of unknown length  
148435 151539: contig of 3105 bp in length  
151540 151559: gap of unknown length  
151560 154468: contig of 2909 bp in length  
154469 154488: gap of unknown length  
154489 156193: contig of 1705 bp in length  
156194 156213: gap of unknown length  
156214 157744: contig of 1531 bp in length  
157745 157764: gap of unknown length  
157765 159461: contig of 1697 bp in length  
159462 159481: gap of unknown length  
159482 160346: contig of 865 bp in length  
160347 160366: gap of unknown length  
160367 161824: contig of 1458 bp in length  
161825 161844: gap of unknown length  
161845 163663: contig of 1819 bp in length  
163664 163683: gap of unknown length  
163684 164031: contig of 348 bp in length  
164032 164051: gap of unknown length  
164052 164700: contig of 649 bp in length  
164701 164720: gap of unknown length  
164721 166088: contig of 1368 bp in length.

## FEATURES

Source  
1. 166088  
/organism="Mus musculus"  
/strain="129/SvEvTACFBr"  
/db\_xref="taxon:10090"  
/chromosome="19"  
/clone="RP22-325J22"  
/sex="male"  
1. 34556  
misc\_feature  
/note="assembly\_name:Contig81"  
34577. 62727  
misc\_feature  
/note="assembly\_name:Contig80"  
62748. 81216

misc\_feature  
/note="assembly\_name:Contig79"  
81237. 92870  
/note="assembly\_name:Contig78"  
92891. 103627  
misc\_feature  
/note="assembly\_name:Contig77"  
103648. 114697  
misc\_feature  
/note="assembly\_name:Contig76"  
114718. 123385  
misc\_feature  
/note="assembly\_name:Contig75"  
123406. 131815  
misc\_feature  
/note="assembly\_name:Contig74"  
131836. 132037  
misc\_feature  
/note="assembly\_name:Contig73"  
132058. 138473  
misc\_feature  
/note="assembly\_name:Contig72"  
138494. 144023  
misc\_feature  
/note="assembly\_name:Contig71"  
144044. 144924  
misc\_feature  
/note="assembly\_name:Contig70"  
144945. 145191  
misc\_feature  
/note="assembly\_name:Contig69"  
145212. 148414  
misc\_feature  
/note="assembly\_name:Contig68"  
148435. 151539  
misc\_feature  
/note="assembly\_name:Contig67"  
151560. 154468  
misc\_feature  
/note="assembly\_name:Contig66"  
154489. 156193  
misc\_feature  
/note="assembly\_name:Contig65"  
156214. 157744  
misc\_feature  
/note="assembly\_name:Contig64"  
157765. 159461  
misc\_feature  
/note="assembly\_name:Contig63"  
159482. 160346  
misc\_feature  
/note="assembly\_name:Contig62"  
160367. 161824  
misc\_feature  
/note="assembly\_name:Contig61"  
161845. 163663  
misc\_feature  
/note="assembly\_name:Contig60"  
163684. 164031  
misc\_feature  
/note="assembly\_name:Contig59"  
164052. 164700  
misc\_feature  
/note="assembly\_name:Contig58"  
164721. 166088  
/note="assembly\_name:Contig57"  
BASE COUNT 44602 a 38872 c 39481 g 42573 t 560 others  
ORIGIN

Query Match 20.9% Score 305.4; DB 2: Length 166088;  
Best Local Similarity 56.3% Pred. No. 2,7e-64;  
Matches 642; Conservative 0; Mismatches 481; Indels 18; Gaps 3;

OY 85 ATGGCATCAAGGAGTTGACGAACCTGCGGACATCGCCCAAGGCGATGAAGAG 144  
DB 11414 ATGGCAATTACGCGCTTGCACAACTAATTCGATGTGGCCCCAGTGCCTCGTGAG 11473  
OY 145 CAGAACTCGAGAGCTACTTCGCCCAAAATGCGCTGCACGACAGATGACATATAC 204  
DB 11474 AATGACATCAAGAGACTTACTTGGTCGCAAGTGCCATCATGCTCCACAGACATCTAC 11533  
OY 205 CAGTTCCTGATTTGATTTGGAAGAGGAGCATGGAACCTCCCAATGATGAGCTGTGA 264  
DB 11534 CAGTTCCTGATTTGCTGCTTCGACAG---GTTGGGAGATGTGTCGAGAAAGAGGTGAG 11590  
OY 265 GTCACATGATTTGCAAGAAATGTTCAACCGCAATTAAGATTACTGCAAGCGGAATC 324  
DB 11591 ACCACACACACCGATGCGGATGTTCACCGTACCTCCGATGATGGAATGCAATC 11650  
OY 325 AAGCCAGTTATGTTTGTGATGCGACAGCTTCCTGATTTGAAGAAACAAGAGCTTGCTAA 384  
DB 11651 AAGCTGTGTACGCTTTGATGAGCAACACACAGCTGAAGTCAAGGACGCTGCGCAAG 11710  
OY 385 AGATTCACAAAAGAGATGATGACCAACCAAGATCTGACTGAGCGATGAGTAGAGAT 444

```

Db 11711 CGCAGTGAAGAGCGCCGCCGAGAGCTGAGAACCACTGCAGCAGGCTCAGAGCTGGGATG 11770
OY 445 AAGATGCGATTGAAAAATTGAGCAGAGAGCTGAAAGTGCACAGAGCAGACAGAGAA 504
Db 11771 GAGGAGGAGGTGGAGAAATTCACCAAGAGGCTGCGTGAAGGTGCACCAAGACACATATGAT 11830
OY 505 GATTGTAACCGGCTATTAAAGACTTATGGGGGTCTCTGTTGAGAGGCACTTCTGAAGCA 564
Db 11831 GAGTCAAAACACCTGCTGAGCCTCATGGGCAATCCCTACTGTATGACACCCAGCAGAGCA 11890
OY 565 GAAGCAGATGTGCAGCCCTTTCATTAACATAGAGTGTGGCTGTCTTGAAGAT 624
Db 11891 GAGGCCAGCTGTGCTCCCTGGCAAGAGCTGTATGCTGCGGCCACGAGAGAC 11950
OY 625 ATGAGCTCCCTACTTCTTGGGGCTCCAGGTTCTCTGCTCAATTAATGATCCAAGTTCC 684
Db 11951 ATGAGCTGCTCACTTGTGGCAGGCCCGGTGAATGCGCACTTAAGTCCAGTAGGCC 12010
OY 685 AAGAAATACCTGTGATGCAATTTGATGTTGCCAAGTTTGGAGGAGCTTGAATCACC 744
Db 12011 AAGAAAGCTCCCATCCAGAGAGTTCATCTGAGCCGCTGCTGAGAGAGCTGGGTCTGAAC 12070
OY 745 ATGAGCACTGCTATGATTTGTCATCCCTGCTGATGATGATGATGATGATGATGATGATGAT 804
Db 12071 CAGGAGCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12130
OY 805 GGTATCGGGGGGAGCAACAGCTCTGAACCTTATTCGTCACATGAGTGGTCCATAGAAAGATC 864
Db 12131 GGGATTGGGCCCAAGCGGGCTGTGGATCTCATCCAGAAACATTAAGAGATCGAGAGATC 12190
OY 865 TTGGAGATCTTAATTAAGACAGATATCAAAATTCCTGAGAGAGCTGGCCTTACCAAGAGCT 924
Db 12191 GTGAGCGGGCTGAGACCCAGCAAGTACCCGTTCCAGAGAGAGCTGCTCCAGAGAGAGCT 12250
OY 925 CGAGCTGTTGTCAGAGAGCTTAATG---TCACATTGATATTCCTGAGAGAGAGAGAGCT 981
Db 12251 CAGCAGCTCTCTCTGAGAGCAGAGATGAGAGCAGAGCTCTGTGAGCTGAAGTGGAGAC 12310
OY 982 GCACCTGATGAGAGAGGCTCATAGTTCTCTGTTAAAGATTAATGTTTCAACAGAGAT 1041
Db 12311 GACCAATTAAGAGAGAGTGTGCAAAATTAATGTTGTTGAAAGAGAGATTTCTGAAGAG 12370
OY 1042 CGGGTGCACAAAGGCCATAGAGAGATCAATCTGCCAAGATTAATCGCCGAGAGAGAA 1101
Db 12371 CGAATTCGAGTGGGTCAAGCGCTGAGTAGAGAGCCGCGAGAGAGAGAGAGAGAGAGAG 12430
OY 1102 CTCGAGTCTTTTTCAGAGCACTGCCACACATCAGACCGCTAAACGGAAGAGAGCT 1161
Db 12431 CTCGATGATTTCTTCAAGGTGACAGCTCACTCTCTCAGC-----TAAAGCGC 12478
OY 1162 TCGGATTAACCAAGCAAGAGAGAGTGGAGAACAAAGAGAGAGAGAGAGAGAGAGAG 1221
Db 12479 AAGGAGCCAGAACCCAGAGGGGCTGCTAAGAGAAAGCAAGAGAGAGAGAGAGAGAG 12538
OY 1222 T 1222
Db 12539 T 12539

```

```

RESULT 12
HSRAD2 1757 bp mRNA linear PRI 07-DEC-1998
DEFINITION Homo sapiens mRNA for flap endonuclease-1.
ACCESSION X76771.1 GI:1905802
VERSION X76771.1
KEYWORDS Ien1 gene; flap endonuclease-1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1757)
AUTHORS Watts, F.

```

```

TITLE Direct Submission
JOURNAL Submitted (16-DEC-1993) F. Watts, University of Sussex, School of
Biological Sciences, Falmer, Brighton BN1 9QG, UK
REFERENCE 2 (bases 1 to 1757)
AUTHORS Murray, J.M., Tavassoli, M., al-Hartley, R., Sheldrick, K.S.,
Lehmann, A.R., Carr, A.M. and Watts, F. Z.
TITLE Structural and functional conservation of the human homolog of the
Schizosaccharomyces pombe rad2 gene, which is required for
chromosome segregation and recovery from DNA damage
JOURNAL Mol. Cell. Biol. 14 (7), 4878-4888 (1994)
MEDLINE 94277093
FEATURES
source Location/Qualifiers
1..1757
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 368..1510
/gene="Ien1"
CDS 368..1510
/gene="Ien1"
/codon_start=1
/product="flap endonuclease-1"
/protein_id="CAA54166.1"
/db_xref="GI:3980293"
/translation="MGIGLAKLIDVAPSAIRENDIKSYERKVAIDASISYFELI
AVROGDDVLQNEGETTSHLMGMFYRTIRMMNGIRKPYVDEGKPKLSBELARSE
RRAEAEKRLQDAAGAEBOEVEKFTKRLVKVTRKHNDCKHLISLMGIPYDAPSEAR
ASCAALVRKAGKYAAATREDMDCLTFGSPVLMHLLTASEAKKLPIDFHLRILOELG
NODRPFVDILIGSDYCESIRIGTGPRAVDLIQKHSEIYRDLDPNKPYPENMLH
KEAHQFLPEVLDPSYSEIRLKSSEPEERELIKFMCGKFOFSERIRSGVRLKLSRQG
STGRDIDDFKFTYGLSLSKRREPERKSTKKKATGAAGKTRKCK"
BASE COUNT 418 a 443 c 542 g 354 t
ORIGIN
Query Match 20.7% Score 302.8; DB 9; Length 1757;
Best Local Similarity 55.0%; Pred. No. 4,5e-64;
Matches 661; Conservative 0; Mismatches 532; Indels 9; Gaps 3;
OY 85 ATGGGCAATCAAGGGTTTGACGAAACTCTGCGGAGCAATGCGCCCAAGCGGATGAAGAG 144
Db 368 ATGGGAATTCAGGCGCTGGCCCAACTAATGCTGATGTGGCCGCCAGTGCCATCCGGAG 427
OY 145 CAGAAATTCGAGAGCTACTTCCGCGGCAAAATGCGCCGCAAGCGGATGAAGAG 204
Db 428 AATGACATCAAGAGCTTCTTTGGCGGTAAGGTGCAATGAGCTTATGACATTTAT 487
OY 205 CAGTCTGATTTGATGTAAGGAGCAGAGCATGGAATCTTCACAAATGAAGCTGTGAA 264
Db 488 CAGTTCCTGATTTGCTGTTGCGCAG---GTGGGAGATGCTGACAGATGAGAGAGGTGAG 544
OY 265 GTCAGTACATTTGCAAGGATGTTACCGCGGACATAGATTACTGGAAGCGGAGATC 324
Db 545 ACCACACAGCCACCTGATGGGATGTTCACCGACCAATTCGATGATGAGAGAGAGAGAG 604
OY 325 AAGCAGTTATGTTTGTGATGGCAAGCTCTCATATGAGAAACAAGAGCTTGTAAA 384
Db 605 AAGCCGTGTATGCTTTGATGCGCAGAGCGCCACAGCTCAAGCTAAGGAGAGGTGCCAA 664
OY 385 AGATACCTCAAAAAGATGATGCAACCAAGATGTGACTGAGGAGTGAAGTGGAGAT 444
Db 665 CGCAGTGAAGCGGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
OY 445 AAGATGCGATTGAAAAATTGAGCAGAGAGCTGAAAGTGCACAGAGCAGACAGAGAA 504
Db 725 GAGCGAGAGGTGGAAAAATTCATAAGCGGCTGTGGAAGGTGCTTAAGCAGACAGAAATGAT 784
OY 505 GATTGTAACCGGCTATTAAAGACTTATGGGGGTCTCTGTTGAGAGGCACTTCTGAAGCA 564
Db 785 GAGTCAAAACATCTCTGAGCCTCATGGGCAATCCCTTATCTGATGACACCAAGTAGAGCA 844
OY 565 GAAGCAGATGTGCAGCCCTTTCATAAAGATTAAGTATGAGTGTGCTGCTTCAAGAGAT 624
Db 845 GAGGCCAGCTGTGCTGCCCTGTGTAAGAGCTGCAAAAGTCTATGCTGCGGCTACGAGAGAC 904

```

```

QY 625 ATGAGCTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGTCAATTAATGATCCAAATTCC 684
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 ATGGACTGCTTCACTTCCGGAGCCCTGTGTAATGACACACCTGACCGCAGTGAAGCC 964
QY 685 AAGAAATACCTGTGATGGAATTTGATGTGCGCAAGGTTTGGAGAGCTTGAACCTCACC 744
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 AAAAAGCTGCAATCGAGGAATTCACCTGAGCCCGATTTCTGCAAGATCGGCTTGAC 1024
QY 745 ATGACCACTTCTTATTTGATTTTGCATCCTGTGTGATGTGACTATTTGATAGATCAAA 804
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1025 CAGGAACAGTTGTGTGATGTGCAATCCTGTAGGAGTACTGTGAGAGATATCGG 1084
QY 805 GGTATCGGGGGGCAACACCTCTGAACCTTATTCCTCAACATGGCTGATGAAGCATC 864
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1085 GGTATTTGGGGCCAAAGGGGCTGTGAGACCTCATCCAGAAAGCAACAGAGCATGAGAGATC 1144
QY 865 TTGAGACATCTTATTAANAAGATATCAATTCCTGAGAGCTGGCTTACCAAGAACT 924
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1145 GTGCGGCGACTTGAGCCCAACAGATACCTGTGCGCAAAATTTGGCTCCACAGAGAGCT 1204
QY 925 CGAGCTTTGTCAGAGAGCTTATG--TCACATTTGATATTTCTGAGCTTAAATGACT 981
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1205 CACAGCTCTTTTGAAGCTGTGAGTGTGACCCAGAGTCTGTGAGCTGAGAGTGAGC 1264
QY 982 GCACCTGATAGAGAGCTCTCATTAAGTTCCGTGTAANAAGATATTTTCAAGAAAT 1041
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1265 GAGCAATGAGAGAGAGCTGATCAATGATGTGTGTGTAAGACAGTTCTGTGAGAG 1324
QY 1042 CGGCTGCAAAAGGCAATAGAGATCAATTCGCCAAGAAATTCGTCGCAAGAGAA 1101
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1325 CGATCCGCGAGTGGGGTCAAGAGGCTGATAGAGACCCGCAAGGACAGCAGGCGCGC 1384
QY 1102 CTGAGAGCTTTTCAAGGCACTGSCAC--CACATCAACGACGCTTAAAGGAGAG 1158
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1385 CTGATGATTTTCTTCAAGGTGACCGGCTCACTCTTCTTCAAGCTAAGCGCAAGACCA 1444
QY 1159 ACTTGCAATTAACAAGCAAGCAAGCTGCGAACAAGAAACAAAGCTGTGGAAGAAG 1218
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1445 CCAAGGATCCACTAAGAAAGGCAAGAGCTGAGGAGCAAGCTTTAAAGAGGGA 1504
QY 1219 AATATCTTGATGCTGTGATGATCACTAGCACTAGCAAGACGAGCTGCGATGATC 1278
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1505 AATATATGTTTCCCATTAATACCTCTTACCCCAAGAAATTTGGCGTCTTGTACCC 1564
QY 1279 TT 1280
    ||
Db 1565 TT 1566
    ||

RESULT 13
BC000323          2031 bp      mRNA      linear      PRI 12-JUL-2001
LOCUS             Homo sapiens, flap structure-specific endonuclease 1, clone
DEFINITION        MCC:8478 IMAGE:2821792, mRNA, complete cds.
ACCESSION         BC000323
VERSION           BC000323.1  GI:12653112
KEYWORDS          MGC.
SOURCE            human.
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         1 (bases 1 to 2031)
AUTHORS           Strausberg, R.
TITLE             Direct Submission
JOURNAL           Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK            NIH-MGC Project URL: http://mgc.nci.nih.gov
CONTACT           MGC help desk
Email:            cgabs@email.nih.gov
Tissue Procurement: DCTD/DBP
cDNA Library Preparation: Rubin Laboratory

```

```

FEATURES
    source
        cdna library Arrayed by: The I.M.A.G.E. Consortium (LINL)
        DNA Sequencing by: National Institutes of Health Intramural
        Sequencing Center (NISC),
        Gaithersburg, Maryland;
        Web site: http://www.nisc.nih.gov/
        Contact: nisc.mgc@nih.gov
        Shevchenko, Y., Weltherby, K.D., Beckstrom-Sternberg, S.M.,
        Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
        Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
        Lim, M., Maduro, O.L., Mastaglio, C., Mastrian, S.D., McCloskey, J.C.,
        McDowell, J., Pearson, R., Snyder, B., Stantirlop, S., Thomas, P.J.,
        Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
        Zhang, L.-H. and Green, E.D.

    CDS
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="MGC:8478 IMAGE:2821792"
        /tissue_type="lung, small cell carcinoma"
        /clone_id="NIH-MGC-7"
        /lab_host="DH10B-R"
        /note="Vector: pOTB7"
        /codon_start=1
        /product="flap structure-specific endonuclease 1"
        /protein_id="AAH00323.1"
        /db_xref="GI:12653113"
        /translation="MTGQGLAKIADVAASAIKRENDIKYRGRKVAIDASLKIOTLI
        AVRGGVLDNERGKTTSLMGKFTRTKMKNGIKPYVFEGRKPOLKSELRKRE
        RRAEAKVLOQAOAGAEDVEKFTRLTKVKNHDEKRLTSLMGIPYDASEAB
        ASCAALVAKKAVYAAATMEDDCLTFGSPVLMHRLVASERKLIPIQEPRLSLDELG
        NOEFDVLCILSDYCESIRGIGPRAVDLQKHRSLEIVRLDPKRYVPVNNWH
        KEAHLFLPEVLEDELKMSPENEELIKFEMGEEKFSEERIRSRGKRLSRSG
        KEAHLDFPFKVTGSLSSAKRKEPEKSGSKTKKAKTGAAGKFRKRGK"

BASE COUNT      533 a      465 c      589 g      444 t
ORIGIN
Query Match      20.7%; Score 302.8; DB 9; Length 2031;
Best Local Similarity 55.0%; Pred. No. 4,7e-64;
Matches 661; Conservative 0; Mismatches 532; Indels 9; Gaps 3;
QY 85 ATGGGATCAAGGCTTTACGAAACTGCTGCGGACATGCGCCCAAGCGGATGAAGAG 144
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 ATGGGATTCAGAGCGCTGCGCAACTATATCTGATGTGGGCCCAAGTCCATCGGAG 198
QY 145 CAGAGTTGAGAGCTACTTGGCGCGCAAAATGCCGCTGACGCGCAGATGACATATAC 204
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 AATGATATCAAGAGCTCTTGGCGGTAAGGTGGCCATATGATGCTTATGACCATTTAT 258
QY 205 CAGTTCTGATGTTGTTGGAAGGACAGCATGGAACCTCTCAACAAATGAAGCTGTGAA 264
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 CAGTTCTGATGTTGTTGTTGCGCAG--GGTGGGATGTGCTGAGATGAGAGGATGAG 315
QY 265 GTTCATGATCTTTCCAGGAATGTTCACCGGACATTAAGATTAAGTGAAGCGGAAATC 324
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 ACCACAGCGCATCTGATGAGGATGTTTACCGCACCATGTGATGATGAAGCGGATC 375
QY 325 AAGCCAGTTATTTTGTATGAGCAAGCCTCGATATGAAGAAACAAGAGCTGTGTA 384
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 AAGCCGCTATGTTTGTATGAGCAAGCGCCACAGCTCAAGTCAAGGAGCGAGCTGCCAAA 435
QY 385 ACATATCTCAAAAAGAGATGAGCAACCAAGATCTGACTGAGAGTGAAGTGAAGAT 444
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 CGCAGTGAAGCGGCGGTGAGGCGAGAGAGCTGACAGAGCTGAGCTGCTGGGGCC 495
QY 445 AAGATGCGATTTGAAAATTTGAGCAAGAGAGCTGTAAGGTCAACAAGCAACAAGAA 504

```



```

* arbitrary. Gaps between the configs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

1	20726:	contig of 20726 bp	in length
20727	20826:	gap of 100 bp	
20827	34665:	contig of 13840 bp	in length
34667	34766:	gap of 100 bp	
34767	46065:	contig of 11295 bp	in length
46066	46165:	gap of 100 bp	
46166	53616:	contig of 7351 bp	in length
53617	53616:	gap of 100 bp	
53617	61065:	contig of 7449 bp	in length
61066	61655:	gap of 100 bp	
61666	67998:	contig of 6833 bp	in length
67999	68098:	gap of 100 bp	
68099	74526:	contig of 6628 bp	in length
74527	74626:	gap of 100 bp	
74627	81299:	contig of 6673 bp	in length
81300	81399:	gap of 100 bp	
81400	8716:	contig of 5917 bp	in length
87317	87416:	gap of 100 bp	
87417	91541:	contig of 4125 bp	in length
91542	91641:	contig of 100 bp	
91642	95731:	contig of 4090 bp	in length
95732	95831:	gap of 100 bp	
95832	101128:	contig of 5297 bp	in length
101129	101228:	gap of 100 bp	
101229	104948:	contig of 3720 bp	in length
104949	105048:	gap of 100 bp	
105049	110306:	contig of 5258 bp	in length
110307	110406:	gap of 100 bp	
110407	112699:	contig of 2263 bp	in length
112697	113069:	gap of 100 bp	
113070	11684:	contig of 3415 bp	in length
116845	116584:	gap of 100 bp	
116585	119599:	contig of 3015 bp	in length
119600	119699:	gap of 100 bp	
119700	121881:	contig of 2182 bp	in length
121882	121981:	gap of 100 bp	
121982	122703:	contig of 722 bp	in length
122704	122803:	gap of 100 bp	
122804	125566:	contig of 2763 bp	in length
125567	125666:	gap of 100 bp	
125667	126645:	contig of 1179 bp	in length
126646	126945:	gap of 100 bp	
126946	129349:	contig of 2404 bp	in length
129350	129445:	gap of 100 bp	
129450	131268:	contig of 1619 bp	in length
131269	131368:	gap of 100 bp	
131369	132565:	contig of 1197 bp	in length
132566	132665:	gap of 100 bp	
132666	134031:	contig of 1366 bp	in length
134032	134131:	gap of 100 bp	
134132	136176:	contig of 2045 bp	in length
136177	136276:	gap of 100 bp	
136277	137418:	contig of 1142 bp	in length
137419	137518:	gap of 100 bp	
137519	138627:	contig of 1109 bp	in length

### FEATURES

Location/Qualifiers

misc\_feature

```
misc_feature      /note="assembly_fragment"
                  20827.  .34666
```

```
misc_feature      /note="assembly_fragment"
                  34767.46065
```

/note="assembly\_fragment"

misc_feature	46166. .53516
misc_feature	/note="assembly_fragment" 53617. .01065
misc_feature	/note="assembly_fragment" 61166. .67998
misc_feature	/note="assembly_fragment" 68099. .74526
misc_feature	/note="assembly_fragment" 74627. .81299
misc_feature	/note="assembly_fragment" 81400. .87316
misc_feature	/note="assembly_fragment" 87417. .91541
misc_feature	/note="assembly_fragment" 91642. .95731
misc_feature	/note="assembly_fragment" 95832. .101128
misc_feature	/note="assembly_fragment" 101229. .104948
misc_feature	/note="assembly_fragment" 105049. .110306
misc_feature	/note="assembly_fragment" 110407. .112969
misc_feature	/note="assembly_fragment" 113070. .116484
misc_feature	/note="assembly_fragment" 116585. .119599
misc_feature	/note="assembly_fragment" 119700. .121881
misc_feature	/note="assembly_fragment" 121982. .122703
misc_feature	/note="assembly_fragment clone_end:T7 vector_side:right" 122804. .125566
misc_feature	/note="assembly_fragment" 125667. .126845
misc_feature	/note="assembly_fragment clone_end:SP6 vector_side:left" 126846. .129349
misc_feature	/note="assembly_fragment" 129450. .131268
misc_feature	/note="assembly_fragment" 131369. .132565
misc_feature	/note="assembly_fragment" 132666. .134031

Query Match	20.78;	Score 302.8;	DB 2;	Length 138627;
Best Local Similarity	55.08;	Pred. No. 1.2e-63;		
Matches 661; Conservative	0;	Mismatches 532;	Indels 9;	Gaps 3

OY	85	ATGGCATCAAGGTTGTGACGAACCGCTGGGGACATCGGCCCAAGGGCATGAAGGAG	144
Db	17332	ATGGGAATTAAGAGCGCTGGCCAAACTAATTGTGATGTGGCCCCCAATGCCATTCGGGAG	17381
OY	145	CAGAAGTTGCAGAGCTACTTGTGGCCGCAAAATCGCCGTGCAGCGCAGCATGAGCATATAC	204
Db	17382	AATGACATCAAGAGCTACTTTGGCCGTAAAGTGGCCATGATGCTGTATGAGCATTTAT	17441
OY	205	CAGTTCCTGATTTATTTGGAAGACGGCATGGAACCTGCACAAATGAAAGTGGTGA	264
Db	17442	CAGTTCCTGATTTCTGTTCGCCAG---GCTGGGATGTGCTGGCAGATGAGAGGTTGAG	17498
OY	265	GTCACTAGTCATTTGCAAGGAATGTTCAACCGGACATTAAGATTACTGGAAGCGGSAATC	324
Db	17499	ACCCACGACCACCTGATGGGCAATGTTACCGCACATTCGCATGAAGGAGAAACGGCATC	17558
OY	325	AAGCATTTATTTTGTATGGCAAGCTCTCTGATATGAGAAACAAAGACTTGGCTAAA	384
Db	17559	AAGCGCTGATTCCTTTGATGGCAAGCGCCACAGCTCAAGTGAAGCGAGCTGGCCAAA	17618
OY	385	AGATACTCAAAAAGACATGATGCAACCAAAAGATCTGACTGAAGGCAATGAGAGTGAAGAT	444
Db	17619	CGCAGTAGCGGGGCTAGGCGAAGAAGCAGCGTACAGCGGCTCAGGCTGTGGGGCC	17678
OY	445	AAAGATGCGATTCAAAATAATTGAGCAAGAGGACTGTAAAGTCAACAGCAACCAACGAA	504





Db 3399 GACGACGAGGAGTGGAAAAATTCACCTAAGCGGCTGTGAAGGTGACTAAGCAGACCAATGAT 3340  
QY 505 GATTGTAAACGGCTATTAAAGCTTATGGGGGTTCCCTGTTAGAGGACCTTCTGAAGCA 564  
Db 3339 GAGTGCAAACATGCTGTAGCCGTCATGGGCATCCCTTAATCTGATGACACCCAGTAGGCA 3280  
QY 565 GAAGCAGAATGTGCACCCCTTTCATAAACGATAAGGTGTGCTGTGCTTCAGAAAGAT 624  
Db 3279 GAGGCCAGCTGTGCTGCTGTGAGGCTGAGCAAGTCTATGCTGCGCTACCGAGAC 3220  
QY 625 ATGAGCTCCCTTACTTTTGGGGCTCCACGGTCTCTTCATTTAATGATCCAAGTTC 684  
Db 3219 ATGAGCTGCCCTACCTTCGAGAGCCCTGTGTAATGCGCACCTGACCTGCCAGTAAAGCC 3160  
QY 685 AAGAAATACCTGTGATTTGATTTGATGTTGCCAAGTTTGGAGGAGCTGAACCTCAC 744  
Db 3159 AAAAAGCTGCAATCCAGGAATTCACCTGAGCCGGAATTCGACAGAGCTGGCCCTGAC 3100  
QY 745 ATGAGCCAGTTCATTTGATTTGTCATCCTGTGTGATGTGACTATTTGATAGCATCAA 804  
Db 3099 CAGGAACAGTTTGTGATCTGTGATCCTGTAGCAGAGTACTGTGAGATATCCGG 3040  
QY 805 GSTATCGGGGGCCCAACACCTCTGAACCTTATTCGTCACATGSGTCCATGAAAAAGCAT 864  
Db 3039 GGTATTGGGCCCAAGCGGGCTGTGGACCTCATCCAGAAAGCAAGAGCATGAGAGATC 2980  
QY 865 TTGGAGAACTTATTAAGAAGATATCAATTCCTGAGGACTGGCTTACCAGAAAGCT 924  
Db 2979 GTGGGGGCACTTGACCCCAACAGTACCTGTGCCAGAAATTTGGCTCCACAGAGAGCT 2920  
QY 925 CGAGGCTTTGTCAGGAGGCTAATG--TCACATTGATATTCCTGAGCTAAATGACT 981  
Db 2919 CACCAGCTCTTCTGAGACTGAGTGTGACCCAGAGCTGTGTGAGCTGAGTGGAGC 2860  
QY 982 GCACCTGATGAGGAGGCTCATAGTTCTGTGTAAGATATGTTTCAAGAGAT 1041  
Db 2859 GAGCCAAATGAAAGAGAGCTGATCAATGATGTGTGAAGAGCAGTCTCTGAGAG 2800  
QY 1042 CGGATGACAAAGGSCATAGAGATCAATTCGCCAAGATTAATGTCGCAAGAGAA 1101  
Db 2799 CGAATCCGCGAGTGGGTGCAAGAGCTGAGTAAGAGCCGCCAAGGAGCAGCAGGCGC 2740  
QY 1102 CTGAGTCTCTTTTCAAGCCACTGCCAC--CACATCAGACCCCTTAAAGCGAAGAG 1158  
Db 2739 CTGATGATTTCTTCAAGGTGACCGGCTCCTCTTCAGTTAGCGCAAGAGGCCAGAA 2680  
QY 1159 ACTTCGATTAACAGAGAGGAGCTGCAAGCAAGAAACAAAGGCTGTGGAAGAAG 1218  
Db 2679 CCGAAGGAGATCCACTAAGAAAGAGACTGGGGCGCAGGAGATTAAAGGGGA 2620  
QY 1219 AAATATCTTGATGCTGTGATGACTAGACTAGCAAGCAGCGGTGGCGTGTATCAG 1278  
Db 2619 AAATAAATGCTTCCCATTAATACCTCTTACCCAGAAATATTGGCGTCTGTATACC 2560  
QY 1279 TT 1280  
Db 2559 TT 2558

Search completed: November 5, 2002, 12:45:26  
Job time : 2515.65 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 08:10:48 : Search time 174.666 Seconds  
(without alignments)  
14380.822 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463  
Sequence: 1 caccagaatagctcgcgccc.....aaaaaaaaaaaaaaaaaa 1463

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N\_Geneseq\_032802.\*  
1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1463	100.0	1463	21	Maize Rad2/FEN-1 c
2	1421.8	97.2	1478	21	Maize Rad2/FEN-1 c
3	1409.8	96.4	1541	21	Maize Rad2/FEN-1 c
4	1350.2	92.3	1381	21	Maize Rad2/FEN-1 c
5	302.6	20.7	1144	20	Human FEN-1 DNA fr
6	302.6	20.7	1144	20	Human FEN-1 CDNA
7	276.6	18.9	2033	20	Human FEN-1 genom
8	276	18.9	1300	23	Drosophila melanog
9	272.2	18.6	1930	20	Mouse FEN-1 CDNA

10	270.2	18.5	1149	20	AAK02109	Yeast FEN-1 CDNA.
11	252.6	17.3	3362	23	ABL14286	Drosophila melanog
12	173.2	11.8	5471	23	ABL10094	Drosophila melanog
13	159.6	10.9	1023	18	AAAT76685	Pyrococcus furiosu
14	159.6	10.9	1023	19	AAAT76685	Pyrococcus furiosu
15	159.6	10.9	1023	19	AAAT76685	Pyrococcus furiosu
16	152.8	10.4	1054	19	AAV53951	Nucleotide sequenc
17	149.6	10.2	1115	19	AAV53984	P. furiosus N-term
18	149.6	9.5	510	22	AAK37479	Human bone marrow
19	139.6	9.5	510	22	AAK37479	Human bone marrow
20	135.6	9.3	1032	20	AAK31850	Probe #8219 for ge
21	134	9.2	34980	22	AAH41224	Pyrococcus heat re
22	128	8.7	514	19	AAV53950	Nucleotide sequenc
23	126.4	8.6	1164	19	AAV53975	Nucleotide sequenc
24	125.6	8.6	1115	19	AAV53978	P. furiosus N-term
25	117.6	8.0	1729	23	AAK37332	DNA encoding novel
26	110.2	7.5	889	19	AAV53974	Nucleotide sequenc
27	98.6	6.7	1011	19	AAV5913	Archaeoglobus fulg
28	98.6	6.7	1011	19	AAV54009	Nucleotide sequenc
29	97.6	6.7	1164	19	AAV53989	M. jansaschii N-te
30	94.8	6.5	386	19	AAV53979	Nucleotide sequenc
31	90	6.2	546	22	ABA63065	Human fetal liver
32	90	6.2	546	22	AAK11490	Human bone marrow
33	90	6.2	546	22	AAK37269	Human brain expres
34	90	6.2	546	22	AAI18098	Probe #8031 for ge
35	89.2	6.1	981	18	AAAT76682	Nucleotide sequenc
36	86.6	5.9	981	19	AAV53948	Nucleotide sequenc
37	86.6	5.9	243	22	AAK50354	Human bone marrow
38	86.6	5.9	243	22	AAI27416	Probe #17349 for g
39	86	5.9	981	19	AAV5837	Methanococcus jann
40	76.2	5.2	777	19	AAV54010	Nucleotide sequenc
41	74.6	5.1	840	19	AAV53977	Nucleotide sequenc
42	69.4	4.7	987	19	AAV54011	Nucleotide sequenc
43	67.8	4.6	296	19	AAV53976	Nucleotide sequenc
44	62.8	4.3	514	19	AAV53985	Nucleotide sequenc
45	62.4	4.3	4551	21	AAV70129	Plasmodium falcipar

#### ALIGNMENTS

RESULT 1	AAK27923	standard; CDNA; 1463 BP.
ID	AAK27923	
XX	AAK27923	
AC	AAK27923	
XX	AAK27923	
DT	12-SEP-2000	(first entry)
XX	12-SEP-2000	
DE	Maize Rad2/FEN-1 CDNA.	
XX	Maize Rad2/FEN-1	
KW	Maize; Rad2/FEN-1; transgenic plant; male sterile plant;	
RW	endonuclease; exonuclease; DNA repair; gene targeting; ss.	
XX		
OS	Zea mays.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	85..1224
XX		/*tag= a
XX		
PN	WO200036109-A1.	
XX		
PD	22-JUN-2000.	
XX		
PF	16-NOV-1999;	99WC-US27147.
XX		
PR	15-DEC-1998;	98US-0112332.
XX		
PA	(PTON-) PIONEER HI-BRED INT INC.	
XX		
PI	Mahaajan FB;	
XX		
DR	WPI; 2000-452026/39.	

DR P-PSDB; AAY95307.  
XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA  
PT recombination and repair in transgenic plants, e.g. for gene targeting  
PT and the production of male sterile plants -  
XX  
PS Claim 1; Page 69-71; 85pp; English.  
XX  
CC The present sequence is that of maize cDNA coding for RAD2/FEN-1  
CC (see AAY95307). The corresponding RNA was isolated from immature  
CC ear tissue from 2 ears of a B73 maize line. Rad2/FEN-1 is a  
CC structure specific endonuclease which under certain conditions also  
CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to  
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.  
CC The protein is involved in the regulation of DNA repair and  
CC recombination in plant systems and therefore may be used for  
CC improving gene targeting during further recombinant DNA protocols  
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential  
CC in DNA replication and nucleotide excision and repair reactions.  
CC The exolytic activity is involved in double strand break repair and  
CC end joining. The protein is also useful in strand exchange  
CC reactions in gene targeting and in the production of male sterile  
CC plants. The efficacy of gene targeting can be improved by the  
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
CC be produced by the down regulation of Rad2/FEN-1 expression.  
XX  
SQ Sequence 1463 BP; 466 A; 292 C; 361 G; 344 T; 0 other;  
Query Match 100.0%; Score 1463; DB 21; Length 1463;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACGAGAAATAGTCGCGCGCGGTTCTTTCGGCCACTCCGGCTCAGCCCGCCGCCA 60  
DB 1 CAGGAGAAATAGTCGCGCGCGGTTCTTTCGGCCACTCCGGCTCAGCCCGCCGCCA 60  
QY 61 CCGCCACAGCCCGCCGACAGAGATGGGATCAAGGGTTGAGCAAACTGTCGGGAC 120  
DB 61 CCGCCACAGCCCGCCGACAGAGATGGGATCAAGGGTTGAGCAAACTGTCGGGAC 120  
QY 121 AATGCGCCCAAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
DB 121 AATGCGCCCAAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
QY 181 GTGCGACCGACATGACATATACAGTTCTGATTTGTTGTTGTTGTTGTTGTTGTTG 240  
DB 181 GTGCGACCGACATGACATATACAGTTCTGATTTGTTGTTGTTGTTGTTGTTG 240  
QY 241 ACTCTCAAAATGAGCTGTGTAAGTACATGATTTGCAAGAAATGTTCAACGGGACA 300  
DB 241 ACTCTCAAAATGAGCTGTGTAAGTACATGATTTGCAAGAAATGTTCAACGGGACA 300  
QY 301 ATAGATTACTGTAAGCGGGAATCAAGCATTATGTTTGTGTTGTTGTTGTTGTTG 360  
DB 301 ATAGATTACTGTAAGCGGGAATCAAGCATTATGTTTGTGTTGTTGTTGTTGTTG 360  
QY 361 ATGAGAAAGAAAGAGCTGTGTAAGTACATGATTTGCAAGAAATGTTCAACGGGACA 420  
DB 361 ATGAGAAAGAAAGAGCTGTGTAAGTACATGATTTGCAAGAAATGTTCAACGGGACA 420  
QY 421 ACTGAGCAGTATGAGTATGAGATTAAGATGATTAAGATTAAGATTAAGATTAAG 480  
DB 421 ACTGAGCAGTATGAGTATGAGATTAAGATGATTAAGATTAAGATTAAGATTAAG 480  
QY 481 AAGGTCAAGAGCAACACAGAGATTGAAAGCGCTATTAGACTTATGGGGTTCCCT 540  
DB 481 AAGGTCAAGAGCAACACAGAGATTGAAAGCGCTATTAGACTTATGGGGTTCCCT 540  
QY 541 GTTGTATAGGACCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
DB 541 GTTGTATAGGACCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600

QY 601 GTTTCGCTGTGTTCTTCAAGATATGAGACTCCCTTACTTTGGGGCTCCAGGTTCCCT 660  
DB 601 GTTTCGCTGTGTTCTTCAAGATATGAGACTCCCTTACTTTGGGGCTCCAGGTTCCCT 660  
QY 661 CCGCATTATATGATGATCAAGTTCAGAGAAATACCTGTGATGAAATTTGATGTTGCCAG 720  
DB 661 CCGCATTATATGATGATCAAGTTCAGAGAAATACCTGTGATGAAATTTGATGTTGCCAG 720  
QY 721 GTTTTGGAGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
DB 721 GTTTTGGAGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
QY 781 TGTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 781 TGTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 841 CAACATGGGTTCATAGAAAGCATCTTGGAATCTTAAATAAAGACAGATATCAATTCCT 900  
DB 841 CAACATGGGTTCATAGAAAGCATCTTGGAATCTTAAATAAAGACAGATATCAATTCCT 900  
QY 901 GAGGACTGGCTTACCAAGAGCTGAGGCTGTTCAAGAGCTTAATGTCATTTGAT 960  
DB 901 GAGGACTGGCTTACCAAGAGCTGAGGCTGTTCAAGAGCTTAATGTCATTTGAT 960  
QY 961 ATTCTGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
DB 961 ATTCTGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
QY 1021 GATTAAGTTCACAGAGATGAGGTTGACAAAGCCATAGAGATGCAATTCCTGCAAG 1080  
DB 1021 GATTAAGTTCACAGAGATGAGGTTGACAAAGCCATAGAGATGCAATTCCTGCAAG 1080  
QY 1081 AATTAATGCTGCAAGAGATGAGGTTGACAAAGCCATAGAGATGCAATTCCTGCAAG 1140  
DB 1081 AATTAATGCTGCAAGAGATGAGGTTGACAAAGCCATAGAGATGCAATTCCTGCAAG 1140  
QY 1141 CCGCTAAACGGAAGGAGACTTGGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
DB 1141 CCGCTAAACGGAAGGAGACTTGGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
QY 1201 AAGCTGTGGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
DB 1201 AAGCTGTGGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
QY 1261 CAGCGGTGCGTATCACTTGTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1320  
DB 1261 CAGCGGTGCGTATCACTTGTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1320  
QY 1321 GGTAAAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
DB 1321 GGTAAAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
QY 1381 TACCAAGTAAACAACTATGCTGTTTACTTCTTGTGCTTGAAGTAAAGTAAAGTAA 1440  
DB 1381 TACCAAGTAAACAACTATGCTGTTTACTTCTTGTGCTTGAAGTAAAGTAAAGTAA 1440  
QY 1441 AAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1463  
DB 1441 AAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1463  
RESULT 2  
AAA27926  
ID AAA27926 standard; cDNA: 1478 BP.  
XX  
AC AAA27926;  
AC  
XX 12-SEP-2000 (first entry)  
XX  
DE Maize Rad2/FEN-1 cDNA.  
XX  
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;  
endonuclease; exonuclease; DNA repair; gene targeting; ss.

XX Zea mays.

XX Key Location/Qualifiers

FH CDS 97..1236

FT /tag= a

XX WO200036109-A1.

XX 22-JUN-2000.

XX 16-NOV-1999: 99WO-US27147.

XX 15-DEC-1998: 98US-0112332.

XX (PION-) PIONEER HT-BRED INT INC.

XX Mahajan PB:

XX WPI: 2000-452026/39.

XX P-PSDB: AAY95310.

PT Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA recombination and repair in transgenic plants, e.g. for gene targeting and the production of male sterile plants

XX Example 1; Page 77-79; 85pp; English.

XX The present sequence is that of maize cDNA coding for RAD2/FEN-1 (see AAY95310). The cDNA was isolated from a library prepared from B73 line seed vitreous endosperm RNA. RAD2/FEN-1 is a structure specific endonuclease which under certain conditions also acts as an exonuclease. RAD2/FEN-1 nucleic acids can be used to produce the RAD2/FEN-1 polypeptides in transgenic plant cells. The protein is involved in the regulation of DNA repair and recombination in plant systems and therefore may be used for improving gene targeting during further recombinant DNA protocols involving plants. RAD2/FEN-1 endonucleolytic activity is essential in DNA replication and nucleotide excision and repair reactions. The exolytic activity is involved in double strand break repair and end joining. The protein is also useful in strand exchange reactions during homologous recombination. These functions may be useful in gene targeting and in the production of male sterile plants. The efficacy of gene targeting can be improved by the overexpression of exogenous RAD2/FEN-1 while male sterile plants can be produced by the down regulation of RAD2/FEN-1 expression.

XX Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other;

XX Query Match 97.2%; Score 1421.8; DB 21; Length 1478;

XX Best Local Similarity 98.5%; Pred. No. 0;

XX Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 7 AATAGCTCGGGGCTTTCTTGGCCATCCGGCTAGCCGCCCGCCACCCGCC 66  
DB 19 AATAGCTCGGGGCTTTCTTGGCCATCCGGCTAGCCGCCCGCCACCCGCC 78  
QY 67 ACAAGCCGCCAGACGATGGCATCAAGGGTTTGACGAACCTCTGCGGACATGCG 126  
DB 79 ACAGCCGCCGACGAGATGGCATCAAGGGTTTGACGAACCTCTGCGGACATGCG 138  
QY 127 CCCAAGGCGATGAAGAGCAAGAGTTCGAGACTCTTGGCGGCAAAATGCGCTGCAC 186  
DB 139 CCCAAGGCGATGAAGAGCAAGAGTTCGAGACTCTTGGCGGCAAAATGCGCTGCAC 198  
QY 187 GCCAGCATGACATATACAGATTCTGATTTGATTTGAAGACAGGAGTGAACCTCTC 246  
DB 199 GCCAGCATGACATATACAGATTCTGATTTGATTTGAAGACAGGAGTGAACCTCTC 258  
QY 247 ACAATGAACTGGTGAAGCACTACTGCAATTTGCAAGGAATGTAACCGGACATTAAGA 306  
DB 259 ACAATGAACTGGTGAAGCACTACTGCAATTTGCAAGGAATGTAACCGGACATTAAGA 318

QY 307 TTACTGAAAGCGGGAATCAAGCCAGTTTATGTTTGGATGGCAAGCTCCTGATATGAAG 366  
DB 319 TTACTGAAAGCGGGAATCAAGCCAGTTTATGTTTGGATGGCAAGCTCCTGATATGAAG 378  
QY 367 AACAGAGCTTGGTAAAGATCTCAAAAAAGAGATGATGCAACCAAGATCTGACTGAG 426  
DB 379 AACAGAGACTTGGTAAAGATCTCAAAAAAGAGATGATGCAACCAAGATCTGACTGAG 438  
QY 427 GCAGTAGAGGTAGAGATTAAGATGCGATTGAAAATTTGACGAAGGAGTGTAAAGTTC 486  
DB 439 GCAGTAGAGGTAGAGATTAAGATGCGATTGAAAATTTGACGAAGGAGTGTAAAGTTC 498  
QY 487 ACAAGCAACACACAGAGGATTTGAAAGCGCTATTAAGACTTATGAGGGGTCTGTTGTA 546  
DB 499 ACAAGCAACACACAGAGGATTTGAAAGCGCTATTAAGACTTATGAGGGGTCTGTTGTA 558  
QY 547 GAGGACCTTTCGAAAGCAAGAGATGTCAGCCCTTTGCAATAACGATTAAGTGTTC 606  
DB 559 GAGGACCTTTCGAAAGCAAGAGATGTCAGCCCTTTGCAATAACGATTAAGTGTTC 618  
QY 607 GCTGTGCTGAGAGATATGACTCCCTTACTTTTGGGCTCCAGGTTCTCTGTCAT 666  
DB 619 GCTGTGCTGAGAGATATGACTCCCTTACTTTTGGGCTCCAGGTTCTCTGTCAT 678  
QY 667 TTAATGATCCAGATTCACAAAGAAATCTGTGATGTAATTTATGTTGGCAAGTTTGG 726  
DB 679 TTAATGATCCAGATTCACAAAGAAATCTGTGATGTAATTTATGTTGGCAAGTTTGG 738  
QY 727 GAGGAGCTTGAATCACCATGAGACAGTTCAATTTGATTCGATCCGTGTGATGTCAC 786  
DB 739 GAGGAGCTTGAATCACCATGAGACAGTTCAATTTGATTCGATCCGTGTGATGTCAC 798  
QY 787 TATTTGATAGCATTAAGGTATCGGGGGCAACACTCTGAAACTTATCTGTAACAT 846  
DB 799 TATTTGATAGCATTAAGGTATCGGGGGCAACACTCTGAAACTTATCTGTAACAT 858  
QY 847 GGCTCCATGAAGAGCATTTTGAGATCTTAATTAAGACAGATCAATTCCTGAGGAC 906  
DB 859 GGCTCCATGAAGAGCATTTTGAGATCTTAATTAAGACAGATCAATTCCTGAGGAC 918  
QY 907 TGCCCTTACCAAGAGCTCGACGCTTGTTCGAAGAGGCTTAATGCAATGATATTCCT 966  
DB 919 TGCCCTTACCAAGAGCTCGACGCTTGTTCGAAGAGGCTTAATGCAATGATATTCCT 978  
QY 967 GAGCTTAATGAGACTGCACTGATGAGAGGGTCTCTTAAGTTTCTGTTAAAGATTAAT 1026  
DB 979 GAGCTTAATGAGACTGCACTGATGAGAGGGTCTCTTAAGTTTCTGTTAAAGATTAAT 1038  
QY 1027 GCTTTCACGAAGATCGGGTGAACAAAGCCATGAGAGAATCAAAATCTGCCAAGATAAA 1086  
DB 1039 GCTTTCACGAAGATCGGGTGAACAAAGCCATGAGAGAATCAAAATCTGCCAAGATAAA 1098  
QY 1087 TCGTGGCAAGGAAGACTGAGTCTTTTTCGAAGCACTGCCACACATATGACACCGCTTA 1146  
DB 1099 TCGTGGCAAGGAAGACTGAGTCTTTTTCGAAGCACTGCCACACATATGACACCGCTTA 1158  
QY 1147 AACAGGAGGAGACTTCGATTAACAGACAGGACACTCGAACAAGAAACAAAGAGCT 1206  
DB 1159 AACAGGAGGAGACTTCGATTAACAGACAGGACACTCGAACAAGAAACAAAGAGCT 1218  
QY 1207 GGTGGAAAGAAAGATTAATCTTGATGCTGTGATGATGATGATGATGATGATGATGATG 1266  
DB 1219 GGTGGAAAGAAAGATTAATCTTGATGCTGTGATGATGATGATGATGATGATGATGATG 1278  
QY 1267 TGCGGTGATCACTTGGCTTAATATTTAACTCCCTGTTTAACTGAGCTTGGTAA 1326  
DB 1279 TGCGGTGATCACTTGGCTTAATATTTAACTCCCTGTTTAACTGAGCTTGGTAA 1338  
QY 1327 AGTTGCTCATGTTTCAGAGCTGGGGTAAAGTAAATGTTGTTGAAGAGATGTTGTTACCA 1386  
DB 1339 AGTTGCTCATGTTTCAGAGCTGGGGTAAAGTAAATGTTGTTGAAGAGATGTTGTTACCA 1398  
QY 1387 GTACAAACCTTATGCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1446



OY	1099	AGACTCGAGTCCCTTTTTCAGGCCACACTGCGACACATCGACACCGCTTAACACGGAAGAG	1158
OY	1099	AGACTCGAGTCCCTTTTTCAGGCCACACTGCGACACATCGACACCGCTTAACACGGAAGAG	1152
OY	1159	ACTTCGGATTAACCAACAGCAGCGAGCTGGCAACAAACAAAGCTGGTGGAAAGAG	1218
Db	1153	ACTTCGGATTAACCAACAGCAGCGAGCTGGCAACAAACAAAGCTGGTGGAAAGAG	1212
OY	1219	AAATATACCTTGAGTGCCTTGATGTGACAACTAAGCACTACGAAACAGCGGTGGGTGATCAG	1278
Db	1213	AAATATACCTTGAGTGCCTTGATGTGACAACTAAGCACTACGAAACAGCGGTGGGTGATCAG	1272
OY	1279	TTTCCTTAGATTAATTTAACTCCCTGTTTAACTCAGAGCTTGGTAAAGTTTGCTCATG	1338
Db	1273	TTTCCTTAGATTAATTTAACTCCCTGTTTAACTCAGAGCTTGGTAAAGTTTGCTCATG	1332
OY	1339	TTTCAAGCTGGGAGTAAGTTGTTGTTGGAAGAGATTGCTATACCAATGAACAACACTT	1398
Db	1333	TTTCAAGCTGGGAGTAAGTTGTTGTTGGAAGAGATTGCTATACCAATGAACAACACTT	1392
OY	1399	ATCGCTGTTTCTTAACTCTCTGTCCTTTGAAGTA 1431	
Db	1393	ATCGCTGTTTCTTAACTCTCTGTCCTTTGAAGTA 1425	
RESULT 4			
AA	AA27925		
XX	AA27925 standard; cDNA: 1381 BP.		
AC	AA27925:		
XX			
DT	12-SEP-2000 (first entry)		
XX			
DE	Maize Rad2/FEN-1 cDNA.		
XX			
KW	Maize: Rad2/FEN-1; transgenic plant; male sterile plant;		
KM	endonuclease; exonuclease; DNA repair; gene targeting; ss.		
XX			
OS	Zea mays.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	37..1176	
FT		/+tag= a	
XX			
PM	WO200036109-A1.		
XX			
PD	22-JUN-2000.		
XX			
PF	16-NOV-1999; 99MO-US27147.		
XX			
PR	15-DEC-1998; 98US-0112332.		
XX			
PA	(PION-) PIONEER HI-BRED INT INC.		
XX			
PI	Mahajan PB;		
XX			
DR	WPI: 2000-452026/39.		
DR	P-PSDB; AAY95309.		
XX			
PT	Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA		
PT	recombination and repair in transgenic plants, e.g. for gene targeting		
PT	and the production of male sterile plants -		
XX			
PS	Example 1: Page 75-76; 85pp; English.		
XX			
CC	The present sequence is that of maize cDNA coding for RAD2/FEN-1		
CC	(see AAY95309). The cDNA was isolated from a cDNA library prepared		
CC	from W23 line tassel (16-18 cm long) POLYA RNA. Rad2/FEN-1 is a		
CC	structure specific endonuclease which under certain conditions also		
CC	acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to		
CC	produce the Rad2/FEN-1 polypeptides in transgenic plant cells.		
CC	The protein is involved in the regulation of DNA repair and		
CC	recombination in plant systems and therefore may be used for		

[illegible]

```

Qy 889 TATCAATTCCTGAGAGCTGGCTTACCAAGAGCTGAGCGCTGTTCAAGAGCCTAAT 948
    |||||||
Db 841 TATCAATTCCTGAGAGCTGGCTTACCAAGAGCTGAGCGCTGTTCAAGAGCCTAAT 900
Qy 949 GTCAATTGATATCTCTGAGCTAAATGACCTGACCTGATGAGAGGGTCTCATACT 1008
    |||||||
Db 901 GTCAATTGATATCTCTGAGCTAAATGACCTGACCTGATGAGAGGGTCTCATAACT 960
Qy 1009 TTCCCTGTAAGATATGTTTCAACGAGATCGGGTGACAAAGGCCATAGAGAATC 1068
    |||||||
Db 961 TTCCCTGTAAGATATGTTTCAACGAGATCGGGTGACAAAGGCCATAGAGAATC 1020
Qy 1069 AAATCTGCCAAGATTAATGCTGCGAAGGAACTGAGTCTTTTTCAGCCCACTGCC 1128
    |||||||
Db 1021 AAATCTGCCAAGATTAATGCTGCGAAGGAACTGAGTCTTTTTCAGCCCACTGCC 1080
Qy 1129 ACCACATCAGACCGCTAAACGAGAGAGACTTCGATTAACAAAGACAGCAGCTGGC 1188
    |||||||
Db 1081 ACCACATCAGACCGCTAAACGAGAGAGACTTCGATTAACAAAGACAGCAGCTGGC 1140
Qy 1189 AACAGAGAAACAAAGCTGCTGGAAGAGAAATTAATTGGATGCTTATGTAACACTA 1248
    |||||||
Db 1141 AACAGAGAAACAAAGCTGCTGGAAGAGAAATTAATTGGATGCTTATGTAACACTA 1200
Qy 1249 CGACTAGAGAAAGCAGCGGTGCGATCACTTCTAGATTAATTAACCTGCTTTTA 1308
    |||||||
Db 1201 CGACTAGAGAAAGCAGCGGTGCGATCACTTCTAGATTAATTAACCTGCTTTTA 1260
Qy 1309 ACTCAGAGCTTGTGTAATAATGTTGCTCATGTTTCAAGCTGGGTAAGTGTGTTTG 1368
    |||||||
Db 1261 ACTCAGAGCTTGTGTAATAATGTTGCTCATGTTTCAAGCTGGGTAAGTGTGTTTG 1320
Qy 1369 AAGAGATTGTGTACCAAGTACAAAACTTATGCTGTTTTT 1411
    |||||||
Db 1321 AAGAGATTGTGTACCAAGTACAAAACTTATGCTGTTTTT 1363

RESULT 5
AA02117
ID AA02117 standard; DNA; 1144 BP.
XX
AC AA02117;
XX
DE 23-APR-1999 (first entry)
XX
DE Human FEN-1 DNA fragment.
XX
KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage; ss.
XX
OS Homo sapiens.
XX
FN US5874283-A.
XX
PD 23-FEB-1999.
XX
PE 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSE/) HSEIH C.
XX
PA (LIEB/) LIEBER M R.
XX
PI Harrington JJ, Hsieh C, Lieber MR;
XX
DR WPI; 1999-179985/15.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing
XX e.g. recombinant polypeptides
XX
PS Claim 3; Column 25; 58pp; English.
XX

```

CC This sequence encodes a human FEN-1 (flap endonuclease) protein. This  
 CC protein can be used in methods for detecting a pathological condition in  
 CC a patient, for diagnostic purposes, for screening for antineoplastic  
 CC agents and carcinogens, for diagnostic staging of neoplasia, for  
 CC producing recombinant flap endonuclease for use as research or  
 CC diagnostic reagents, for producing transgenic nonhuman animals expressing the  
 CC polypeptides, for producing transgenic nonhuman animals expressing the  
 CC novel polypeptides encoded by a transgene. The invention also provides  
 CC novel molecular cloning techniques and reagents involving cleavage of  
 CC a flap or nick with a flap endonuclease.

SO Sequence 1144 BP; 292 A; 283 C; 353 G; 216 T; 0 other.

Query Match 20.7%; Score 302.6; DB 20; Length 1144;

Best Local Similarity 56.6%; Pred. No. 1.2e-66;

Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

```

Qy 85 ATGGGCAATTCAGAGGTTTGAAGAACTGCGGAGCAATAGCGCCCAAGCGCATGAAGAG 144
    |||||
Db 1 ATGGGAATTCAGAGGCTTGGCCAAACTAATTGCTGATGCGCCCAAGCGCATCGGGAG 60
Qy 145 CAGAGATTGAGAGCTACTTGGCCGCAAAATCGCCGTGACGCGCAGCATATATAC 204
    |||||
Db 61 AATGACATCAAGAGCTACTTGGCCGTAAGGTGCGCATGATGCTTATGAGCATTTAT 120
Qy 205 CAGTCTGATTTGATGTTGGAAGAGACAGCATGGAACCTCACAATAAGCTGGTGA 264
    |||||
Db 121 CAGTCTGATTTGATGTTGGAAGAGACAGCATGGAACCTCACAATAAGAGGTGAG 177
Qy 265 GTCACTAGTCAATTTGCAAGAAATGTTCAACCGAGATATAGATTACTGGAAGCGGAATC 324
    |||||
Db 178 ACCACCGCCACCTGATGAGGCTATGTTCTACCCACCATCTCGATATGAGAGAGGCTATC 237
Qy 325 AAGCGATTATGTTTGTATGAGCAAGCCCTGATGATGAAACAAACAGACCTGCTTAA 384
    |||||
Db 238 AAGCCCTGATGTTCTTTGATGAGCAAGCCGCAAGCTCAAGCTCAAGTCAAGAGTGGCCAA 297
Qy 385 AGATACTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGCGAGTAGAGGTAGAGAT 444
    |||||
Db 298 CGCAGTGAAGCGGCGGCTGAGCAGAGAGAGCTGACAGAGCTGACAGGCTGCTGGGCGC 357
Qy 445 AAGATGCGGATGAAAAATTTGAGCAAGAGACGCTGTAAGGTCACAAAGCAACACAGCA 504
    |||||
Db 358 GAGCAGAGGTGAAAAATTTCACTAGCGGCTGTTGAGAGGTACTTACAGACAGCAATGAT 417
Qy 505 GATTTGAACGGCTATTAGACTTATGAGGGGTTCCTGTTGTAGAGGCACTTCTGAAGCA 564
    |||||
Db 418 GAGTGCAAACATCTGCTGAGGCTCATGCGCATTCCTTATCTTGAGCACCCAGTAGGCA 477
Qy 565 GAAGCAAGATGTCAGCCCTTGGCAATAAGAGATAGGTTGCGTGTGTCGTAAGAGAT 624
    |||||
Db 478 GAGGCAAGCTGTGCTGCTGCTGGTGAAGGCTGCAAAAGTATATGCTGCGCTACCAAGGAC 537
Qy 625 ATGACATCCCTTACTTTTGGGCTCCACGCTTCCTTGCATTTAATGATCAAGTTC 684
    |||||
Db 538 ATGACATGCTTCACTTGGCAGCCCTGTGTAATGTCAGACCTGACAGTGAAGACC 597
Qy 685 AAGAAATTAACCTGTGATGGAATTTGATTTGCCAAGGTTTGGAGACCTGAGACTCAC 744
    |||||
Db 598 AAAAGCTGCCAATTCAGGAATTCACCTGAGCCGGATTTCTCAGAGAGCTGGCCGTGAC 657
Qy 745 ATGACAGATTCAATGATTGTCATCTGTGTGATGATGATGATTTGATAGATCAAA 804
    |||||
Db 658 CAGGACAGTTTGTGATGTCATCTGTAGGAGTATGATGATGATGATGATGATGATG 717
Qy 805 GGTATCGGGGGGCAAAACAGCTTGAACATTTATTCGTCAACATGGGTCCATAGAAAGATC 864
    |||||
Db 718 GGTATGCGGCCCAAGCGGCTGTGACATTCAGAAACCAAGAGCATGAGAGATC 777
Qy 865 TTGAGAGATTTAATTAAGACAGATTAATAATCTGAGAGATGCGCTTACAGAAAGCT 924
    |||||
Db 778 GTGCGGGAATTCACCCCAACAAAGTACCTGTGTCAGAAATAATGGCTCCACAGAGAGCT 837

```

Query Match	20.7%	Score 302.6	DB 20	Length 1144
Y	925	CGAGCGTGTGTTCAAGAGCCCTAATG---TCACATGTGGATATTCCTGACGTAATAATGACT	981	
D	838	CACACAGCTCTTCTTGAACACTGAGGTGTGAGACCCAGAGCTGTGTGAGCTGAAGTGTGAGC	897	
Y	982	GCACCTGATGAGGAGGCTCTCATAGTTCCTGTGTAAAGATTAATGTTCAACGAGAT	1041	
D	898	GAGGCAATGTAAGAGAGCTGATCATCAAGTTTCATGTGTGTAAAGCATGTTCTGTGAGGAG	957	
Y	1042	CGGCTGACAAAGGCCATGAGAGATCAATATTCGCCAAGATTAATGTCGCAAGGAGA	1101	
D	958	CGATCCGCGAGTGGGGGTCAAGAGGTGTAGTAAAGCCGCCAAGCAGCAACCCAGGGCCGC	1017	
Y	1102	CTGAGATTCCTTTTTCAGCCCACTGCCACCATCAGCACC	1142	
D	1018	CTGGATGATTTCTTCAAGGTGACCGGCTCATCTCTTCACG	1058	
RESULT 6				
AA02107		standard; cDNA; 1144 BP.		
AC	AA02107:			
XX	23-APR-1999	(first entry)		
DT	Human FEN-1 cDNA.			
XX				
DE				
XX				
KW	FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;			
XX	neoplasma; antineoplastic agent; cleavage; ss.			
OS	Homo sapiens.			
FT	Key	Location/Qualifiers		
FT	CDS	1..1144		
FT		/*tag= a		
FT		/product= "FEN-1"		
XX	US5874283-A.			
XX	23-FEB-1999.			
PD				
XX	30-MAY-1995;	95US-0455968.		
PF				
XX	30-MAY-1995;	95US-0455968.		
PR				
XX	(HARR/) HARRINGTON J J.			
PA	(HSIE/) HSIEH C.			
PA	(LIEB/) LIEBER M R.			
XX				
PI	Harrington JJ, Hsieh C, Lieber MR;			
XX				
DR	WPI: 1999-179985/15.			
DR	P-PSDB: AAW92504.			
XX				
XX	DNA encoding flap endonuclease polypeptides - useful for producing			
PT	e.g. recombinant polypeptides			
XX				
PS	Disclosure; Fig 1B; 58pp; English.			
XX				
CC	This sequence encodes a human FEN-1 (flap endonuclease) protein. This			
CC	protein can be used in methods for detecting a pathological condition in			
CC	a patient, for diagnostic purposes, for screening for antineoplastic			
CC	agents and carcinogens, for diagnostic staging of neoplasia, for			
CC	producing recombinant flap endonuclease for use as research or			
CC	diagnostic reagents, for producing antibodies reactive with the novel			
CC	polypeptides, for producing transgenic nonhuman animals expressing the			
CC	novel polypeptides encoded by a transgene. The invention also provides			
CC	novel molecular cloning techniques and reagents involving cleavage of			
CC	a flap or nick with a flap endonuclease.			
XX				
50	Sequence 1144 BP; 292 A; 283 C; 353 G; 216 T; 0 other;			

[illegible]



Db 1018 CTGCATGATTTCTTCAAGGTGACCGGCTCACTCTCTTACG 1058

RESULT 7  
ID AAX02111 standard; DNA; 2033 BP.  
AC AAX02111;  
XX  
XX 23-APR-1999 (first entry)  
XX  
XX Human FEN-1 genomic DNA.  
XX  
XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;  
XX  
XX neoplasma; antineoplastic agent; cleavage; ss.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 104..1240  
XX FT /\*tag= a  
XX FT /product= "FEN-1"  
XX  
XX US5874283-A.  
XX  
XX 23-FEB-1999.  
XX  
XX 30-MAY-1995; 95US-0455968.  
XX  
XX 30-MAY-1995; 95US-0455968.  
XX  
XX (HARR/) HARRINGTON J J.  
XX (HSIE/) HSIEH C.  
XX (LIEB/) LIEBER M R.  
XX  
XX Harrington JJ, Hsieh C, Lieber MR;  
XX  
XX WPI; 1999-179985/15.  
XX P-PSDB; AAW92508.  
XX  
XX DNA encoding flap endonuclease polypeptides - useful for producing  
XX e.g. recombinant polypeptides  
XX  
XX Disclosure; Fig 5A-B; 58pp; English.  
XX  
XX This sequence encodes a human FEN-1 (flap endonuclease) protein. This  
XX protein can be used in methods for detecting a pathological condition in  
XX a patient, for diagnostic purposes, for screening for antineoplastic  
XX agents and carcinogens, for diagnostic staging of neoplasma, for  
XX producing recombinant flap endonuclease for use as research or  
XX diagnostic reagents, for producing transgenic nonhuman animals expressing the  
XX novel polypeptides encoded by a transgene. The invention also provides  
XX novel molecular cloning techniques and reagents involving cleavage of  
XX a flap or nick with a flap endonuclease.  
XX  
XX Sequence 2033 BP; 500 A; 493 C; 579 G; 461 T; 0 other:  
SQ

Query Match 18.9%; Score 276.6; DB 20; Length 2033;  
Best Local Similarity 55.9%; Pred. No. 66-60;  
Matches 638; Conservative 0; Mismatches 479; Indels 24; Gaps 5;

QY 85 ATGGGCAATCAAGGTTTGAGAAACTCTGCGGACATGCGCCCAAGCGATGAAGAG 144  
DB 104 ATGGAAATTCACGGCTTGCCAACTAATGCTGATGTGCGCCCATGCGCATCCGTGAG 163  
QY 145 CAGAAGTTTCAGAGCTACTTCGCGCCCAAAATCGCGTCAGCCAGCATGACATATAC 204  
DB 164 AATGACATCAAGAGCTACTTGTGCGCAAGTGCGCATGCGTCATGACATCTAC 223  
QY 205 CAGTTCCTGATGTAGTTGGAAGACAGCGATGGAACCTCACAATGAGCTGGTGA 264  
DB 224 CAGTTCCTGATGTGCTGCTGTCTAG--GTTGGGATGTGCTGAGAGAGAGGCTGAG 280

QY 265 GTCACTAGTCAATTTGCAAGCAATGTTCAACCGGACAATAAGATFACGAAAGCGGAATC 324  
DB 281 ACCACCAGCC---TGATGGGCAATGTTCTACCGTACCATGCG---CATGCAATATGGCATC 334  
QY 325 AAGCCAGTTTATGATTTTATGAGCAAGCCTCTGATATGAGAAACAGAGCTGCTAAA 384  
DB 335 AAGCTGTGATGCTCTTTTATGAGCAAGCCTCTGATATGAGAAACAGAGCTGCTAAA 394  
QY 385 AGATACTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGGCGATGAGGTAGCATAT 444  
DB 395 CGCAGTGAAGAGCGCGCCGAGGCTGAGAGCAACTGCAAGGCTCAGCAGGCTGGATG 454  
QY 445 AAGATGCGCATTTAAATAATTGAGCAGAGAGACTGTAAGGTCACAAGGCAACACAGNA 504  
DB 455 GAGGAGAGGTGAGAAAGTTCCACCAAGAGGCTGCTGATGAGTACCAAGCAACATATGAT 514  
QY 505 GATTGTAACGCGCTATTAAAGACTTATGGGGTTCCTGTTGTAGAGGACCTTCTGACGA 564  
DB 515 GAGTGCACAAACCTGCGAGAGCCCTCATGGGATCCCTTACCTTGATGACCCAGCGAGCA 574  
QY 565 GAAGCAGATGTGACAGCCCTTTGCAATACGATTAAGTGTGCTGCTTCCAGAGAT 624  
DB 575 GAGGCCAGCTGTGCTGCTGCGCCGCAAGGCTGCAAGGTATGCTGCGCCACGGAAGAC 634  
QY 625 ATGGACATCCCTTACTTTTGGGCTCCACGAGTTCCTTCGTCATTTAAATGATCCAACTTC 684  
DB 635 ATGGACTGCTCTACTTTTGGCAGCCCGCTGCTAATGAGACACTTAACCTGACAGTGAAGC 694  
QY 685 AAGAAATACCTGTATGATGAATTTGATGTTGCCAAGTTTGGAGAGCTGTAACCTACC 744  
DB 695 AAGAAGCTGCCATCCAAAGAGTTCCATCTGAGCCGCTCTGAGAGAGGCTGCTGAAC 754  
QY 745 ATGACCAATTCATGATTTGATGATCTGATGATGATGATGATGATGATGATGATGATGAT 804  
DB 755 CAGGAGCAATTTGATGATCTGATGATCTGATGATCTGATGATGATGATGATGATGATGAT 814  
QY 805 GGTATCGGGGGGCAACAGCTCTGAACTTAATTCGTCACATGAGGCTCATAGAAGCATC 864  
DB 815 GGCATTTGGCGGCAAGCGGGCTGTGATCTCATCCAGAAACATAGAAGCATGAGGATC 874  
QY 865 TTGGAAGATTTTAATTAAGACAGATATCAAAATTCCTGAGGACGCGCTTCCAAAGCT 924  
DB 875 GTGAGCGGCTGGACCCCGACAGATCCCGTTCAGAGAACTGCGTCCCAAGAAAGCC 934  
QY 925 GCAGCCTTTGTCAGAGAGCTTAA--TGTCACATTTGATATTTCTGAGCTTAATGAGACT 981  
DB 935 CAGCAGCTCTTCTGAGACCGAAGTACTGACCCAGAGCTGTGAGAGCTGAGAG 994  
QY 982 GCACCTGATGAGAGGCTCTCATAGTTTCTGTAAGATAATGCTTCAACGAAGAT 1041  
DB 995 GAGCCAAATGAAGAAGATGTCATAATTTATGTGTGTGTAAGAGCAATTTTGAAGAG 1054  
QY 1042 CGGGTGAACAAAGGCTATGAAAGATCAAAATTCGCCAAGAAATAATGTCGAGAGAGA 1101  
DB 1055 CGAATTCGAGTGGGGTGAACGGGTGATGTAAGAGCCGCCAGGCGACACCCAGAGAGCC 1114  
QY 1102 CTCGAGTCTTTTCAAGCAACTGCGCACATCAGACCGCTAAAGCGAGAGAGACT 1161  
DB 1115 CTCGATGATTTTCTCAAGTGAAGGCTCACCTCTCTCAGC-----TAAAGCG 1162  
QY 1162 TCGGATTAACCAAGACAGGCTGCGAACAAGAAACAAAGGCTGAGGAAACAAGAA 1221  
DB 1163 AAGGAGCCAGAACCAGAGGGGCTGCTAAAGAAAGAAAGACTGGGGGAGCGGGAGAG 1222  
QY 1222 T 1222  
DB 1223 T 1223

RESULT 8  
ABLI4287  
ID ABLI4287 standard; cDNA; 1300 BP.



```

XX ABL14287;
AC
XX
XX 26-MAR-2002 (first entry)
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 37343.
XX
XX Drosophila: developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI: 2001-656860/75.
XX
XX P-PSDB: ABB70184.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 37343; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB857737-AB872072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 1300 BP; 346 A; 313 C; 369 G; 272 T; 0 other;
SQ
Query Match 18.9%; Score 276; DB 23; Length 1300;
Best Local Similarity 55.2%; Pred. NO. 7e-60;
Matches 581; Conservative 0; Mismatches 465; Indels 6; Gaps 2;
QY 85 ATGGGCATCAAGGGTTGACGAACCTGTCGCGACATGCGCCCAAGCGCATGAAGAG 144
DB 1 ATGGGAATTTGGGCTTATCTAAGCTCATTTGCGCATCTGCGCGACAGCGGATTCGGA 60
QY 145 CAGAGTTCGAGAGCTACTTGGCGCGAAATCGCGTTCAGACGCGACATGACATATAC 204
DB 61 AGTGAGATGAAGCAATTTTTCGTGCGAAGTACCAATGCATGTACTAATGCTCTGTA 120
QY 205 CAGTTCCTGATTGTAGTTGGAAGAGCAGCATGGAACCTGCACAAATGAAGCTGTGA 264
DB 121 CAGTTCCTGATTGCGCGCTCGCGAAGGC--GCCAGTTGGCGACCGTAAATGGTGA 177
QY 265 GTCACTATGCTATTGCAAGAAATGTTCAACCGGACAAATTAAGATTACGGAAGCGG 324
DB 178 CCCAGCTCCCTTAATGAGCATGTCTACCGCACATCCGATTGCTGCAACGGAATC 237
QY 325 AAGCCAGTTTATGTTTGTATGAGCAAGCCCTCTGATATGAAGAAACAGAGCTTGCTAA 384
DB 238 AAGCCGATATATGCTTGTGAGCAAGCAACACATCTCAAGTCCGCTGAGCTGGCGA 297
QY 385 AGATACTCAAAAAGAGATGATGACCAACCAAGATCTGACTGAGCGATGAGAGTGA 444

```

```

DB 298 CGGCCGAGCGCGGAGGAAGCGAGAGGCACTGTAAGGCGGCCACCGATGGGAGAT 357
QY 445 AAGATGCGATTGAAAAATGAGCAAGACGACTGTAAAGTCCAGAGCAACCAACGAA 504
DB 358 GATGCGGATATGCAAAAGTTTAAATCGCGATGTGTCCGGGTAAGAGAGCGCGCAAA 417
QY 505 GATTGTAAACCGCTATTAGACTTATGAGGGGTTCTGTTAGAGGCACTTCTGAAGCA 564
DB 418 GAGGCGCAAGAACTGCTACACTAATGAGTGGTCCCATGTATGTATGACACCGTCAAG 477
QY 565 GAAGCAGAAATGTGACGCCCTTTGATTAAGCATAAAGGTGTCTCTGTTCTGACAAAT 624
DB 478 GAGGCCAGATGTGGCGCTGTGTGAAGCGTAAGGTTATGTCACCCCGACGAGGAT 537
QY 625 ATGACATCCCTTACTTTTGGGGGCTCACGGTCTCTGTCATTAAAGATCCAAAGTTC 684
DB 538 ATGATGCGCTCACATTCGATGCGATCAAAACTGTGAGTACTTACCTAACCGAGGCA 597
QY 685 AAGAAATACCTGTGATGAAATTTGATGTCACAAAGTTTGGAGAGCTTGAATCACC 744
DB 598 CGAAAGATGCCCGTCAAGGAGTTTCAGCTACGACAAAGCTGTGGAAAGTCTGGCCATTAA 657
QY 745 ATGACACAGTTCAATGATTTGTGATCCTGTGTGATGTGACTATTGTATGATCAAA 804
DB 658 AATCGAGAGTTCAATGATCTATGATTTCTGTGCGATTGATGATGAGAGCATCAAG 717
QY 805 GGTATCGGGGGGCAACACGCTGCAAACTTATGTCACATGAGTGTCAATGAAGCATC 864
DB 718 GGTATGAGACCCCAACGAGCGATGCAACTGATCAACCTATGCGGATATGAGACTATT 777
QY 865 TTGGAGATCTTAAATTAAGACATATCAATTCCTGAGGACTGGGCTTACCAAGAGCT 924
DB 778 CTGATTAACCTGAGCTAGCAATATACCGGTGCGGAGACTGGAACATCAAGGTGCG 837
QY 925 CGAGCTTGTTCAGAGAGCTTAATGTCAC--ATTGATATTCCTGAGCTAAATGAGACT 981
DB 838 CGGGAACCTTCAATCAAGCAACCGAGGTAGCTGATGCCACTCCATAGATCTCAATG 897
QY 982 GCACCTGATGAGGGGGTCTCATAGTTCTCTGTGTAAGATTAATGTTGATCAAGAA 1041
DB 898 GAGCGGATGAGAGGCGCTTGTCAAGTTCTGCGGCGACCGGACATTCACGAGAG 957
QY 1042 CGGTGCAAAAGCCATAGACAGATCAAAATGCGCAAGATTAATGTCGCAAGAGA 1101
DB 958 CGCGTTCGCAACGCTGCGCAAAATCAACCTGATCAAAATCCAGCAGGCCAGACTC 1017
QY 1102 CTGAGTCTTTTTCAGAGCCCACTGCCACGAC 1133
DB 1018 CTGATAGCTTTTAAAGACACTGCCACGAC 1049

```

RESULT 9  
 ID AAX02108 standard; cDNA; 1930 BP.  
 XX  
 AC AAX02108;  
 XX  
 DT 23-Apr-1999 (first entry)  
 XX  
 DE Mouse FEN-1 cDNA.  
 XX  
 KW FEN-1; mouse; flap endonuclease; cleavage; ss.  
 KW neoplasma; antineoplastic agent; cleavage; ss.  
 OS Mus sp.  
 XX  
 PN US5874283-A.  
 XX  
 PD 23-FEB-1999.  
 XX  
 PF 30-MAY-1995; 95US-0455968.  
 XX

```

PR 30-MAY-1995; 9505-0455968.
XX (HARR/) HARRINGTON J J.
PA (HSIEH/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX Harrington JJ, Hsieh C, Lieber MR:
PI MPI: 1999-179985/15.
DR P-PSDB; AAW92505.
XX
XX DNA encoding flap endonuclease polypeptides - useful for producing
PT e.g. recombinant polypeptides
XX
XX Disclosure: Fig 2B; 58pp; English.
XX
XX This sequence encodes a mouse FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel
CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.
XX
SQ Sequence 1930 BP; 482 A; 459 C; 553 G; 436 T; 0 other:
Query Match 18.6%; Score 272.2; DB 20; Length 1930;
Best Local Similarity 55.1%; Pred. No. 7.6e-59;
Matches 629; Conservative 0; Mismatches 488; Indels 24; Gaps 4;
QY 85 ATGGGCATCAAGGGTTTGAAGAACTGCTGGCGACATGCGCCCAAGGCATGAAGAG 144
DB 1 ATGGGAATTCACGGCTTCCCAAACTAATGCTGATGTGGCCGCCCAAGTCCCATCGTAG 60
QY 145 CAGAAATTTGAGAGACTACTTGGCGCCCAAAATGCGCGTGCAGCCAGTCATATAC 204
DB 61 AATGACATCAAGAGCTACTTGTGCTAATGAGCCATGATGATGCTTCCATGACATCAG 120
QY 205 CAGTTCCTGATTTGATTTGAGAGAGCAGCATGAAACTCTCAAAATGAAAGCTGTGAA 264
DB 121 CATTCTCTGATTTCTGTCTGCTGAG---GGTGGGATGTGCTCAGAAACGAGAGCGTAG 177
QY 265 GTCACTAGTCATTTGCAAGCAATGTTCAACCGACATTAAGATTACTGGAACGGGAATC 324
DB 178 ACCACCAAGCCTGATGGCATGTTAT-----GCCAAACCATCCGATGGAGAATGGCATC 231
QY 325 AAGCCAGTTTATTTTGTGATGCAAGCCTCCTGATATGAAAGAAACAAAGCTTGCTAAA 384
DB 232 AAGCTCTGTACTCTTTTATGTCAAACCAACCAACGACTGAAGTCAGGCGAGCTGGCAAG 291
QY 385 AGATACTCAAAAAGAGATGTCAGCAACCAAGATCTGATGAGGAGCAGTAGAGTAGAT 444
DB 292 CGCAGTGAGAGCGCGCGCGCTGAGAAAGCAACTGCACAGCCTCAGAGAGCGCTGGAGT 351
QY 445 AAGATGCGGATTTGAAAAATGAGCAAGAGAGACTGTAAGAGTCCAGAGCAACAGCAAGAA 504
DB 352 GAGGAGAGAGGTGAGAGTTTCAACAGAGAGCTGTCAGAGTCCACAGCAACCAATGAT 411
QY 505 GATTGTAAAGGCTATTAACTTATAGGGGTTCTGTTGTAGAGGCACCTTCTGAAGA 564
DB 412 GAGTGCAAAACCTCGTGAACCTTCATGGGATCCTTACTTGTATGACACCGAGCGAGCA 471
QY 565 GAAGCAGATGTGACAGCCCTTTGCAATAAGATGAAGTGTGCTGTGCTTCAAGAAAT 624
DB 472 GAGGCCAGCTGTGCTGCTGGCAAGGCTGCAAAAGTCTATGCTCGGCGACAGGAGAC 531
QY 625 ATGAGCTCCCTTACTTTTGGGGCTCCAGCGTCTCTTCTGATTTAATGATCCAAATTC 684
DB 532 ATGAGCTGCTCCTACTTTTGGACAGCCCGTGTCTAATGCGACACTTAACCTGCACTGAGGCC 591

```

```

QY 685 AAGAAATACTGCTGATGGAATTTGATTTGGCCAAAGCTTTTGGAGAGCTTGAACCTCACC 744
DB 592 AAGAACTGCGCATTCACAAAGTTTCATCTGAGCCCGCTCTCAGAGAGATGGGTGTAAG 651
QY 745 ATGACAGCTTCAATTTGATTTGTCATCTGTGTGATGTGATCTATTGTGATCAAA 804
DB 652 CAGGACAGATTTGTGATCTGTGATCTGTGATCTGTGATCTGTGATCTGTGATCTGTGAT 711
QY 805 GGTATGGGGGGGCAACAGCTCTGAAACTTATTCGCAACATGGGCTCAATAGAAAGATC 864
DB 712 GGCATTTGGGCCCAAGCGGCTGTGATCTCATCAGAAACATTAAGATCAAGAGATC 771
QY 865 TTGAGACATCTTAATTAAGACAGATATCAAAATTCCTGAGAGCTGGCTTACCAAGAGCT 924
DB 772 GTGAGCGGCTGCAACCCCAAGTACCCCGTTCCAGAACTGCTCCACAAGAGAGCC 831
QY 925 CGACGCTTTTGAAGAGCCTAA---TGTACATTTGATTTCTGTAGCTAAATGAGACT 981
DB 832 CAGCAGCTCTTCTGAGAGCCAGAGTATGAGACCCAGAGTCTGTGAGAGCTGAAGTGAAGC 891
QY 982 GCACCTGATGAGAGGGTCTCATAGTTTCTGTAAAGATATGTTTCAACGAAAT 1041
DB 892 GAGCCAAATGAAGAGAGTTGTCATATTTATGTGTGTAAGAAAGCAATTTCTGAGAG 951
QY 1042 CGGGTGACAAAGGCCATAGAGATCAATCTGCCAAGATTAATCTGTGCAAGAGAGA 1101
DB 952 CGAATTCGAGTGGGGTCAAGCGGCTGAGTAAAGAGCCGCGCAGGAGCAGCAGGAGAGCCG 1011
QY 1102 CTGAGATCTTTTTCAGGCACTGTCACATCAAGCAGCCGCTTAAGAGAGAGACT 1161
DB 1012 CTCATATTTTTCAGAGGTGACAGGCTCCTCTCTCAGC-----TAAAGCCG 1059
QY 1162 TCGGATTAACAAAGCAGGAGCTGCAACAGAAACAAAGAGCTGTGTAAGAGAGAA 1221
DB 1060 AAGAGCGCAAGACCAAGGGGCTCTTAAGAAAGCAAAAGATGCGGAGAGCGGGAAG 1119
QY 1222 T 1222
DB 1120 T 1120

```

RESULT 10  
AA02109  
ID AA02109 standard; cdna; 1149 BP.  
XX  
XX AA02109;  
AC  
DT 23-APR-1999 (first entry)  
XX  
DE Yeast FEN-1 cDNA.  
XX  
XX FEN-1; yeast; flap endonuclease; detection; diagnosis; carcinogen;  
KW neoplasia; antineoplastic agent; cleavage; ss.  
XX  
OS Saccharomyces cerevisiae.  
XX  
XX US5874283-A.  
PN  
PD 23-FEB-1999.  
XX  
PE 30-MAY-1995; 9505-0455968.  
XX  
PR 30-MAY-1995; 9505-0455968.  
XX  
PA (HARR/) HARRINGTON J J.  
PA (HSIEH/) HSIEH C.  
PA (LIEB/) LIEBER M R.  
XX  
PI Harrington JJ, Hsieh C, Lieber MR:  
XX  
XX MPI: 1999-179985/15.  
DR  
DR P-PSDB; AAW92506.  
XX

PT DNA encoding flap endonuclease polypeptides - useful for producing  
 PT e.g. recombinant polypeptides  
 PS Disclosure; Fig 3B; 58pp; English.  
 XX  
 CC This sequence encodes a yeast FEN-1 (flap endonuclease) protein. This  
 CC protein is used in a method to isolate novel human FEN-1 proteins for  
 CC detecting a pathological condition in a patient, for diagnostic purposes,  
 CC for screening for antineoplastic agents and carcinogens, for diagnostic  
 CC staging of neoplasia, for producing recombinant flap endonuclease for use  
 CC as research or diagnostic reagents, for producing antibodies reactive  
 CC with the novel polypeptides, for producing transgenic nonhuman animals  
 CC expressing the novel polypeptides encoded by a transgene. The invention  
 CC also provides a novel molecular cloning techniques and reagents involving  
 CC cleavage of a flap or nick with a flap endonuclease.  
 CC  
 XX  
 XX  
 SQ Sequence 1149 BP; 410 A; 182 C; 279 G; 278 T; 0 other;

Query Match 18.5%; Score 270.2; DB 20; Length 1149;  
 Best Local Similarity 55.5%; Pred. No. 2e-58;  
 Matches 599; Conservative 0; Mismatches 453; Indels 27; Gaps 3;

85 ATGGGATCAGAGGTTTGAAGAACTGCGCGGACATGCGCCCAAGGATGAAGAG 144  
 1 ATGGGATTAAGAGTTTGAATGCAATTAATGCGAACAATGTTCTCTGCTATGAGAAA 60  
 145 CAGAGATTGAGAGCTACTTCTGCGCCCAAAATCGCCGTCGACGACATGATATATAC 204  
 61 AGGATATCAGAGCTTTTGGCAGAAAAGTTGCCATGATGCTCTATGCTATAT 120  
 205 CAGTTCTCATTTGTAGTTGGAAGAGCAGATGAAACTCTACAAATGAAGCTGTGAA 264  
 121 CAGTTTATTTAGTCTGTAAGACAGCAAGAGCGTGGCAGTTGACCAATGAACGCGTGAA 180  
 265 GTTCATGATCTTTGCAAGAAATGTTCAACCGGACATTAAGTTACTGGAGAGGGAATC 324  
 181 ACAAGCTCAGACTGTGATGATGTTTATAGACACTGGAATGATGTAAGGATATC 240  
 325 AAGCCAGTTTATGTTTGTATGAGCAAGCCTCTGATATGAGAAAGAGAGCTGTGTA 384  
 241 AAGCCTTTGATGTTGTCAGCGCAAACTCCAGCTTTGAATCTGATGATGAGCAAG 300  
 385 AGATCTCAAAAAAGAGATATGCAACCAAGATCTGACTGAGGCACTAGAGTGAAGAT 444  
 301 CGGTCTTCAAGAGAGGTGGAACAGAAAAAACTGCGAGAGCAACACAGATTTGAA 360  
 445 AAGATGCGATGAAAAATTTAGCAAGAGACTGTAAGGTCAAGGCAAGCAACAGAA 504  
 361 AGATGACAGCAAGAAAGA-----AGATGTTGAAGGTCTCAAAAGGCAATATGAA 411  
 505 GATTGTAACGCGCTATTAAGACTTATGAGGGGTCTGTTGTAGAGGCACCTTCTGAAGA 564  
 412 GAAGCCCAAAAAATTAAGTACTAGCAATATGGAATCCCATATATATAGCGCAAGAGCT 471  
 565 GAAGCAGATGTGCAAGCCTTTGCTAATAACGATAGGTGTGCTGTGCTTCAAGAT 624  
 472 GAGGCTCAATGTGCTAGTTGGCAAGAGGAAAGGTATGCTGATGCCACAGATGAAGAT 531  
 625 ATGAGACTCCCTACTTTTGGGGCTCCAGCGTTCCTCGCATTTAATGATCCAAATTC 684  
 532 ATGAGACACTCTGTTATAGAACACCCCTCTTGTGAGCAATTTGACTTTTCAAGGCC 591  
 685 AAGAAATATCCTGTATGAGAAATTTGATGTTGCCAAGTTTGGAGAGGCTTGAACAC 744  
 592 AAGAGAGAACCGCATTCACGAATATAGTACTGAATGTTTGGAGAGACTCGACTTGACA 651  
 745 ATGAGCAATTTATGATTTGTCATCTGTGTGATGTCATTTGATGATGATGATCAAA 804  
 652 ATAGAGCAATTTGATGTTGCTAATAATGCTGTTGATGATGCTGATGATGATGATG 711  
 805 GGTATCGGGGCAACAGCTGGAATCTTATGTCATGAGGAGGCTTATGATGATGATG 864  
 712 GGTGTTGGTCAAGTACAGCCTTAAATGATTAATAAAGCATGATGATTCATCAAAAAATC 771

QY 865 TTGAGAACTT-----AATAAGACAGATATCAAAATTCCTGAGACTG 909  
 DB 772 GTGAGATTATTAATCTGGGAGATCAAAACATTAATGAAATCCAGAAAGACTGG 831  
 QY 910 CCTTACAGAAAGCTCGAGCTGTTTCAGAGAGCCTTAATGCAATTTGATA--TTCT 966  
 DB 832 CCTTACAAACAGAAAGTGTGTTCTTGACCTGAAGTTATGATGTAAGCAATA 891  
 QY 967 GACTTAATATGACTGACTGATGAGAGAGGTCATTAAGTTTCCGTGTAAGATAT 1026  
 DB 892 AACTTGAAATGTGCGCCACCAAGAGAAAGAACTTATCGAGATTTATGATGATAG 951  
 QY 1027 GATTTCAGCAAGATCGGTCAGCAAGGCCATGAGAAAGATCAATCTGCCAAGATTA 1086  
 DB 952 AATTCAGTGAAGAAAGATTAATCTGTTATATCAAGATTTGAAAAAGCTTGAATCT 1011  
 QY 1087 TCGTCGCAAGAACTCGAGTCTTTTTCAGCAAGCCATGCAACACAGCAGCAGCT 1145  
 DB 1012 GGCATTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1070

RESULT 11  
 ID ABL14286  
 ID ABL14286 standard; cDNA; 3362 BP.  
 XX  
 AC ABL14286;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 37340.  
 XX  
 KW Drosophila: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 OS  
 OS Drosophila melanogaster.  
 PN  
 PN WO200171042-A2.  
 PD  
 PD 27-SEP-2001.  
 XX  
 PE 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 DR MPI; 2001-656860/75.  
 DR P-PDB; ABB70183.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 PS Claim 1: SEQ ID NO 37340; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-AB72072).  
 The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pt\_sequences.  
 Sequence 3362 BP; 979 A; 709 C; 750 G; 924 T; 0 other;

Query Match	17.3%;	Score 252.6;	DB 23;	Length 3362;
Best Local Similarity	55.2%;	Pred. No. 8.8e-54;		
Matches 536;	Conservative 0;	Mismatches 429;	Indels 6;	Gaps 2;
QY	166	GGCCGCAAAATCGCGTCGACGGCCAGCATATACCATGTTCTGATTTAGTTGA	225	
DB	1144	GGTCGCAAGTACATCGATGCTAGTAGTGCTGTACAGTTCCATCCGCGGCCG	1203	
QY	226	AGACAGGATGGAACCTCACAATGAAAGCTGTGAATCTACTACTGTTGCAAGGA	285	
DB	1204	TCGGAAGGC---GCCAGTTGGCCACCGTAATAGTGATCCACGTCCTTAATGGGC	1260	
QY	286	ATGTTCAACCGACAAATTAAGTACTGGAAGCGGAATCAAGCCAGTTATGTTTGTAT	345	
DB	1261	ATGTTCTACCGCAATCCGATTTGCTGGAACAAGATCAAGCCGGATATGCTTCGAC	1320	
QY	346	GGCAGCCTCTGATATGAAGAAACAAGCTTGCTAAAGACTCTCAAAAGAGATGAT	405	
DB	1321	GGAAAGCCACAGATCTCAAGTCGGTGAGCTGGCGAAGCGCCGAGCGCGAGGAA	1380	
QY	406	GCAACCAAGATCTGACTGAGCGAGTAGAGATGAAGATGCGATTGGAATAATG	465	
DB	1381	GCGGAGAGGACTGAAAGGCGCCACCGATGCGGAGATATGCGGTATCGAAAGTTT	1440	
QY	466	AGCAGAGACTGTAAAGTACACAGGCAACAGACAGATTTTAAACGCTATTAGA	525	
DB	1441	AATCGCGATTTGGTCCGGGTAAAGAGAGAGAGCCAAAGAGCCAAAGAACTGCTACA	1500	
QY	526	CTTATGGGGGTTCTGTTTAGAGGACCTTCTGAAAGCAGACAGATATGCAACCCCTT	585	
DB	1501	CTATATGGTGTGCGCTATGTTGATGACCCGTGGAGACGAGGCGCCGCTGCGGCTGTG	1560	
QY	586	TGATTAAGCATGAGTGTGCTGCTTCAAGATGTGACCTCCCTTACTTTTGGG	645	
DB	1561	GTCAAAGCTGGAAGGTTTATGCCACCGCCACGAGAGATTTGGATGCCCTTCACATTGGA	1620	
QY	646	GCTCCAGGTTCTCTGCTCATTTAATGATCAAGTTCAGAAATTAACCTGATGAA	705	
DB	1621	TTTACAAACTGTTGAGATACCTTACTACAGGAGCAGAAAGATGCCGTCACAGAG	1680	
QY	706	TTTGATGTTCCCAAGGTTTGGAGAGCTTGAACCTCACCATGACCGCTCATTTGATTG	765	
DB	1681	TTTACGTCACACAAGCTGTGGGAAGGCTGCGCATTTAACAATGAGAGTTTCATTGATCTA	1740	
QY	766	TGATCCTGTGTGATGTGACTATTGTGATGACATCAAGGTATCGGGGGCAAAACGCT	825	
DB	1741	TGATTTCTGCTGGTTGCCATTTACTGTGAGAGCATCAAGGATTTGACCCCAAGCGG	1800	
QY	826	CTGAAACTTATTCGTCAACATGGGTCCATAGAAAGCATTTGGAGAATCTTAATAAGAC	885	
DB	1801	ATCGAACTGTATCAACACTATCGGGATATAGAGACTATTCGTGATTAACCTGACTAGC	1860	
QY	886	AGATATCAATTCCTGAGGAGTGGCTTACCAAGAAAGCTGAGCCCTGTTCAGAGACCT	945	
DB	1861	AAATACACCGTGGCCGGAAGTGAAGTGAAGTGGCGGGGAACCTCTTCATGGAACCG	1920	
QY	946	AATGTCAC---ATTGATATTCTGAGCTAAATGAGCTGAGTGAAGAGGCTCTC	1002	
DB	1921	GAGGTACTGATGGCGCATTCATAGATCTCAATGGGTGAGCGCGATGAGAGGCGCTT	1980	
QY	1003	ATTAAGTTCTCTGGTAAAGATATATGTTTCAACGAGATGGGTGACAAAGCCATAGAG	1062	
DB	1981	GTCAAAGTTTCTGCGCGCGGCAAGCTCAACGAGAGGCGCTTCCACAGCGTGCCAAA	2040	
QY	1063	AAGATCAATCTGCGCAAGATTAATGTCGCAAGAGAGCTCGAGTCTTTTCAAGCCA	1122	
DB	2041	AAGGTATGAATCAAGCAGGCGCCAGACTAGGTGAGTCTGATGCTTTTAAGACA	2100	
QY	1123	ACTGCCACAC 1133		
DB	2101	CTGCCGACGAC 2111		

RESULT 12			
ABL10094			
ID ABL10094 standard; cDNA: 5471 BP.			
XX			
AC ABL10094;			
XX			
DT 26-MAR-2002 (first entry)			
XX			
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 24764.			
XX			
KW Drosophila; developmental biology; cell signalling; insecticide;			
KW pharmaceutical; gene; ss.			
XX			
OS Drosophila melanogaster.			
XX			
PN WO200171042-A2.			
XX			
PD 27-SEP-2001.			
XX			
PF 23-MAR-2001; 2001WO-US09231.			
XX			
PR 23-MAR-2000; 2000US-191637P.			
XX			
PR 11-JUL-2000; 2000US-0614150.			
XX			
PA (PEKE ) PE CORP NY.			
XX			
PI Venter JC, Adams M, Li PMD, Myers EW;			
XX			
DR WPI; 2001-656860/75.			
XX			
DR P-PSDB; ABB65991.			
XX			
PT New isolated nucleic acid detection reagent for detecting 1000 or more			
XX			
PT genes from Drosophila and for elucidating cell signalling and cell-cell			
XX			
PT interactions -			
XX			
PS Claim 1: SEQ ID NO 24764; 21pp + Sequence Listing; English.			
XX			
CC The invention relates to an isolated nucleic acid detection reagent			
XX			
CC capable of detecting 1000 or more genes from Drosophila. The invention is			
XX			
CC useful in developmental biology and in elucidating cell signalling and			
XX			
CC cell-cell interactions in higher eukaryotes for the development of			
XX			
CC insecticides, therapeutics and pharmaceutical drugs. The invention			
XX			
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA			
XX			
CC sequences (ABL161840-ABL16175) and the encoded proteins			
XX			
CC (ABB57737-ABB72072).			
XX			
CC The sequence data for this patent did not form part of the printed			
XX			
CC specification, but was obtained in electronic format directly from WIPO			
XX			
CC at ftp.wipo.int/pub/published_pcl_sequences.			
XX			
SQ Sequence 5471 BP; 1422 A; 1223 C; 1325 G; 1501 T; 0 other;			
XX			
Query Match	11.8%; Score 173.2; DB 23; Length 5471;		
Best Local Similarity	54.0%; Pred. No. 1.4e-33;		
Matches 377; Conservative	0; Mismatches 318; Indels 3; Gaps 1;		
QY	439	GGAATTAAGATGCGATTGAATAATGAGCAAGAGACTGTAAAGTCAACAGCAAC	498
DB	4	GGAATGATGCGGTTGCGAAGATTTAATCGCCATTGGTCCGGGTAAAGAGAGAC	63
QY	499	AACGAATTTGAAACGCTATTAAAGACTTATGGGGTCTCTGTTAGAGGACCTTCT	558
DB	64	GCCAAAGAGGCCAAGGAAGTGTCTGACACTAATGGGTGCTGCTATGTTGATGACCGTGC	123
QY	559	GAAGCAGAAAGCATATGTCAGCCCTTTGATTAACGATTAAGGTGTTGCTGTTCA	618
DB	124	GAAGCGAGAGCCCGAGTGTGCCCTCTGTGTAAGCTGGAAGATTATGCGACCGCAGC	183
QY	619	GAAGATATGAGTCCCTTCTTTTGGGCTCCACGTTTCTCTGCTATTTAATGATCA	678
DB	184	GAGATATGATGCCCTTCACATTGCGATTAACAAACTGTTAGATACCTTACTACAGC	243
QY	679	AGTTCCAAAGAAATACCTGTGATGGAATTTGATGTTGCCCAAGTTTGGAGAGCTTGA	738

```
Db 244 GAGGACGAAAGATGCCCTCAAGAGATTCAGCTACGACAAAGCTGTTGAAAGTCTGGCC 303
QY 739 CTCACCATGACACCATTCATTTGATTTGCAACCTGATGATGATGATGATGATGATGATG 798
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 ATTAACATGACAGATTCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 363
QY 799 ATCAAGATATCGGGGGGCAACAGCTGTAACCTTATTCGTCAACATGATGGTTCATAGAA 858
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 364 ATCAAGGGTATTTGACCCCAAGCGAGATGCACTGATCAACATATGCGGATATAGAG 423
QY 859 AGCATCTGGAGATCTTAAATAAGACATATCAATTCCTGAGACTGGCTTGACCA 918
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 ACTATTTGATATACCGGCTCTAGCAAAATGACCGTGCCGCAAGACTGGAATCAAG 483
QY 919 GAAGCTGACGCTTTGTTCAAGAGACCTAATGTCAATTTGATTTGATTTGCT--GAGCTA 975
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 484 GTGGCGGGGAACTCTTCAATCGAAGCGAGATGCTGATGCTGATGCTGATGCTGATGCTGATG 543
QY 976 TGGACTGACCTATATAGAGAGGCTTCATTAAGTTTCTGTTAAAGATTAATGTTTCAAC 1035
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 TGGGTGAGCGGATGAGAGGGGCTTGTCAAGTTTCTGCTGCGCGGACCGGAGTTCAAC 603
QY 1036 GAAGATCGGGTGACAAAGGCGCATAGAGATCAATTCGCCAAGATTAATTCGTGCAA 1095
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 604 GAAGAGCGGCTTGCAGACGCTGCCAAAAGCTGATGAATAATCCAGCAGGCCAGACTCAG 663
QY 1096 GGAAGACTCGAGTCTTTTTCAGCCCACTGCCACCAC 1133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 GTGAGACTCGATAGCTTCTTTAAGACATGCCACGAC 701
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
AAT76685
ID AAT76685 standard; DNA; 1023 BP.
XX
AC AAT76685;
XX
DT 14-APR-1998 (first entry)
XX
DE Pyrococcus furiosus FEN-1 endonuclease gene coding sequence.
XX
KW Nucleic acid cleavage; 5' nuclease; DNA cleavage; RNA cleavage;
KM Invader directed cleavage; FEN-1; endonuclease; ds.
XX
OS Pyrococcus furiosus.
XX
PN WO9727214-A1.
XX
PD 31-JUL-1997.
XX
PF 22-JAN-1997; 97WO-US01072.
XX
PR 02-DEC-1996; 96US-0759038.
PR 24-JAN-1996; 96US-059491.
PR 12-JUL-1996; 96US-0682853.
PR 29-NOV-1996; 96US-0756386.
PR 02-DEC-1996; 96US-0758314.
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichev VI;
PI Olive DM, Prudent JR;
XX
DR WPI; 1997-393613/36.
XX
P-PSDB; AAM24216.
XX
PT Thermostable structure-specific nuclease(s) - used for detection and
PT characterisation of nucleic acid sequences and variations in nucleic
XX acid sequences
XX
PS Example 28; Page 283-285; 457pp; English.
XX
CC This sequence comprises the coding region of the gene encoding
```

```
CC Pyrococcus furiosus (Pfu) FEN-1 endonuclease (see AAM24216).
CC It was obtained by PCR amplification (see AAT76682-83). Large
CC scale preparation of recombinant Pfu FEN-1 was performed using
CC E. coli as host. Pfu FEN-1 is a thermostable enzyme. It can be
CC used in novel methods for the detection and characterisation of
CC nucleic acid sequences and variations in nucleic acid sequences.
SQ Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other;
Query Match 10.9%; Score 159.6; DB 18; Length 1023;
Best Local Similarity 54.6%; Pred. No. 1.9e-30;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;
QY 134 CGATTAAGAGGACGAGACTTCAGAGACTTCTTGCGCCCAAAATGCGGTGAGCGCCAGCA 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 29 CAAGAAAGAAATTTGATTAAGAAACCTATACGGGAAAAAATATGCAATCGACCTCTTA 88
QY 194 TGACATATACCAAGTCTCTGATTTGATTTGAGAGGACGCGCATGGAAACCTCAAAATG 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 89 ATGCAATCTACCAATTTTGTCCACAAATTAAGACAAAGATGAACTCCACTTATGATTT 148
QY 254 AAGCTGTGAAGTCACTAGTCAATTTGCAAGAAATGTTCAACCGGACAAATAGATTACTGG 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 CAAGGGTATGATTAACCTCCACCTAAGCGGGCTCTTTTACAGGACAAATAAAGCTAATGG 208
QY 314 AAGCGGAAATCAAGCCACTTATGTTTGTATGGCAAGCTTCGTGATATGAGAAACAAG 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 AGCTGGATTAACCTGTGTATGTTTGTATGAGAAACCTCCAGAAATTCAAAAGAAAG 268
QY 374 AGCTTGTAAAGATATCTCAAAAGAGATGATGCAACCAAGATCTGATGAGGACAGTAG 433
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 269 AGCTGAAAAAAGAAAGAACGAGAGAGAGAGCTGAAGAAATGAGAGAAAGACACTTG 328
QY 434 AGGTATGAGATTAAGATGCGATTGAAAAATTTAGCAAGAGAGCTGTAAAGGTCAACAAGC 493
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 AAAAAGGAGATAGAGAGAAAGAAAAATATGCCAAAGACCAACGAGGTAAATGAA 388
QY 494 AACACAAAGAAATTTGAACGCTATTAAGACTTATGAGGGGTTCCGTGTAGAGGAC 553
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 TGTCTATGAGAGATGCAAAAAAAGCTTTAGAGCTTATGAGGAAATCTTATAGTTCAAGC 448
QY 554 CTTTGAAGCAGAGCAAGATGTGACGCCCTTGTGCATAAAGATAGGTGTGCTGTG 613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 CTAGGAGGAGAGGCCCAAGCTCATATATAGGCCGCAAAAGGAGGAGGTATATGATCGG 508
QY 614 CTTCAGAAATATGAGACTCCCTTAATTTTGGGCTCCAGCGTTCCTTCATTTAATGG 673
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 CTAGTCAAGATTAGATTCCCTACTTTTGGAGCTCCAAAGACTGTGAAGAACTTAACA 568
QY 674 ATCCAGTTCCAGAAATATCCTGTGATGATTTGATTTG 715
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 TAACAGGAAAAAGAAAGTTGCTGGAAAAATGTTAGCTGG 610
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
AAV65840
ID AAV65840 standard; DNA; 1023 BP.
XX
AC AAV65840;
XX
DT 02-FEB-1999 (first entry)
XX
DE Pyrococcus furiosus FEN-1 endonuclease gene ORF.
XX
KW Nucleic acid detection; multiple sequential invasive cleavage;
KM FEN-1; endonuclease; nuclease; ds.
XX
OS Pyrococcus furiosus.
XX
PN WO9842873-A1.
XX
PD 01-OCT-1998.
XX
```

XX	24-MAR-1998;	98WO-US05809.
XX	24-MAR-1997;	97US-0823516.
XX	(THIR-) THIRD WAVE TECHNOLOGIES INC.	
XX	Brow MAD, Hall JG, Kwiatkowski RW, Lyamichev VI;	
XX	Mast AL, Vavra SH;	
XX	WPI: 1998-557036/47.	
XX	P-PSDB: AAM79970.	
XX	Detecting target nucleic acid by sequence-specific cleavage of	
XX	complex with two specific oligonucleotides - used to detect	
XX	cytomegalovirus DNA	
XX	Example 2db; Page 316-317; 524pp; English.	
XX	This is the nucleotide sequence of the open reading frame encoding	
XX	the FEN-1 endonuclease (see AAM79970) of <i>Pyrococcus furiosus</i> (Pfu).	
XX	It was obtained from genomic DNA by PCR amplification (see	
XX	also AAM65838-39). The PCR product has been ligated into vector	
XX	pUC99A, and FEN-1 was expressed in <i>E. coli</i> cells. The invention	
XX	relates to means for the detection and characterisation of nucleic	
XX	acid sequences, and variations in nucleic acid sequences. It also	
XX	relates to methods for forming a nucleic acid cleavage structure on	
XX	a target sequence and cleaving this structure in a site-specific	
XX	manner, preferably using a thermostable structure-specific nuclease	
XX	such as FEN-1. Cleavage of the cleavage structure by the	
XX	nuclease indicates the presence of specific nucleic acid sequences	
XX	or specific variants. The invention further relates to methods for	
XX	the separation of nucleic acid molecules based on charge, methods	
XX	for the detection of non-target cleavage products via the formation	
XX	of a complete and activated protein binding region, and methods for	
XX	the detection of nucleic acid from various viruses (e.g. human	
XX	cytomegalovirus) in a sample. The method amplifies the detection	
XX	molecule rather than the target itself, is less subject to	
XX	contamination than exponential amplification processes, and allows	
XX	many targets to be analysed in a single reaction.	
XX	Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other;	
XX	Query Match	10.9%; Score 159.6; DB 19; Length 1023;
XX	Best Local Similarity	54.6%; Pred. No. 1.9e-30;
XX	Matches 318; Conservative	0; Mismatches 264; Indels 0; Gaps
XX	134 CGATGAGAGAGCAGAAAGTTGCGAGCTACTTGGCCGCAAAATCGCGTCGACGCCAGCA	193
XX	29 CAGCAAAAGCAAAATTTGATTAGAAAACCTATTACGSGGAAAAATTCGAAATCGAGCTCTTA	88
XX	194 TGACGATATACCGATCCGATTCGATTTGTAGTTGGAGAGCAGCATGAAATCCATCAAAATG	253
XX	89 ATGCATCTACGCAATTTTGTGTCACATATAGACAGAAAGTGGAACTCCACTTATATGATT	148
XX	254 AACCTGTGTAAGTCTAGTATCATTTGCCAAGGAATGTTCAACCGGACATTAAGATTACTGG	313
XX	149 CAAAGGTTAAATTAACCTCCACCTTAAGCGGCGCTTTTACAGACACATAAATCACTAATGG	208
XX	314 AACGGGAATCAAGCCAGTTATGTTTGTATGGCAAGCTCCTGATATGAAAGAAACAAAC	373
XX	209 AGCTGTGAATTAACCTGTATGTTTGTATGATGAGAAACCTCCAGATTTCAAAAATAAAG	268
XX	374 AGCTGTATAAAGATTAATCAAAAAGAGATATGCAACCAAGATCTGACTGAGGACAGTAG	433
XX	269 AGCTCAAAAAAAGAGGAAGCAGAGAGAGAACTGAAAGAAAGTGGAGAGAGCACTTGG	328
XX	434 AGGTAGAGATTAAGATGCGATTGAAAATTTGACCAAGAGAGACTGTAAAGTTCACAAAGC	493
XX	329 AAAAAGAGAGATAGAGAGAGCAAGAAAATATGCCCCAAAAGCAACCGAGGTAAATGAAA	388
XX	494 AACCAACCAAGATTTAAACGGCTATTAGACTTTATGGGCGTTCCTGTTGTAGAGGAC	553
XX	389 TGCTCATCGAGAGTCAAAAAAACTCTTTAAGCTTTAGGGAATTTCTATATGTTCAAGCAC	448

QY	554	CTTCTGAAGCAGAAAGCAGAAATGTCAGCCCTTTCCATTAACGATTAAGTGTGCTGTG	613
Db	449	CTAGCGAGGAGGAGGCCCAAGCTGATTAATGAGCCGCGAAGGAGAGCGCTGTATGATCGG	508
QY	614	CTTCAGAGAAATGACATCCCTTACTTTTGGGGCTCCACGGTTCCTTCGTCAATTAATGG	673
Db	509	CTAGTCAGATTTACGATTCCTACTTTTGGAGGCTCCAGACTGTGTAGAACTTAACAA	568
QY	674	ATCCAGTTCCAGAAATFACCTGTGATGGAATTTGATGTTG	715
Db	569	TACACGAAAAAGAAAGTTCCTGGGAAAAATGCTACGTG	610
RESULT 15			
AAV53951	ID AAV53951 standard; DNA: 1023 BP.		
AAV53951;			
AC	21-DEC-1998	(first entry)	
XX			
XX	Nucleotide sequence of the Pfu FEN-1 endonuclease.		
XX			
XX	Pfu FEN-1 endonuclease; Tag gene; structure-specific nuclease;		
KM	mutant; DNA polymerase; bacteria; fungi; protozoa; RNA virus;		
KM	hepatitis C virus; HCV; thermostable; ds.		
XX			
OS	Pyrococcus furiosus.		
XX			
XX	Key	Location/Qualifiers	
FT	CDS	1..1023	
FT		/*tag= a	
FT		/product= "Pfu FEN-1 endonuclease"	
XX			
XX	W09823774-A1.		
PN	04-JUN-1998.		
XX			
PF	26-NOV-1997;	97MO-US21783.	
XX			
XX	02-DEC-1996;	96US-0758314.	
PR	29-NOV-1996;	96US-0757653.	
XX			
PA	(THIR-) THIRD WAVE TECHNOLOGIES INC.		
XX			
PI	Kaiser MW, Lyamichev VI, Lyamicheva N;		
XX			
DR	WPI; 1998-322748/28.		
XX	P-PSDB; AAW59940.		
PT	Thermostable structure-specific nuclease(s) derived from mutant DNA		
PT	polymerase(s) - useful for detecting mutant allele(s) or strains of		
PT	microorganisms		
XX			
PS	Example 29; Pages 278-279; 472pp; English.		
XX			
XX	This is the nucleotide sequence encoding the Pfu FEN-1 endonuclease,		
CC	used in the method of the invention. In this process thermostable		
CC	structure-specific nucleases are derived from mutant DNA polymerases,		
CC	which can be used for detecting mutant alleles or strains of		
CC	microorganisms. The structure-specific nucleases can be used in		
CC	mixtures, compositions and kits to treat nucleic acid, e.g. for		
CC	detection of wild type and mutant alleles of genes, for detection		
CC	and/or identification of strains of microorganisms such as bacteria,		
CC	fungi, protozoa, especially for detection of RNA viruses such as the		
CC	hepatitis C virus (HCV).		
XX			
SQ	Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other;		
Query Match 10.9%; Score 159.6; DB 19; Length 1023;			
Best Local Similarity 54.6%; Pred. No. 1,9e-30;			
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;			









Db 301 ATAAAGTACTGTAGAGCGGGAATCAAGCCAGTTATGTTTTGATGCGCAAGCCTCCTGAT 360  
QY 361 ATGAAGAACACAGAGCTTCTAAAAATACTCAAAAAGAGATGATGCAACCAAGATCTG 420  
Db 361 ATGAAGAACACAGAGCTTCTAAAAATACTCAAAAAGAGATGATGCAACCAAGATCTG 420  
QY 421 ACTGAGGAGTGTAGTAGAGATTAAGATGAGTGAATAAATGAGCAGAGAGAGCTGTA 480  
Db 421 ACTGAGGAGTGTAGTAGAGATTAAGATGAGTGAATAAATGAGCAGAGAGAGCTGTA 480  
QY 481 AAGGTCAACAGCAACAGCAAGATGTAACGCGCTATTAAAGACTTATGAGGCGTTCTT 540  
Db 481 AAGGTCAACAGCAACAGCAAGATGTAACGCGCTATTAAAGACTTATGAGGCGTTCTT 540  
QY 541 GTTGATGAGGACCTTCTGAAGCAGAGAGATGATGAGCAGCCTTTCGATTAAGAGTAAG 600  
Db 541 GTTGATGAGGACCTTCTGAAGCAGAGAGATGATGAGCAGCCTTTCGATTAAGAGTAAG 600  
QY 601 GTGTCGCTGCTTCTCAGAGATATGACCTCCCTTACTTTTGGGCGTCGACGGTTCCTT 660  
Db 601 GTGTCGCTGCTTCTCAGAGATATGACCTCCCTTACTTTTGGGCGTCGACGGTTCCTT 660  
QY 661 CGTCAATTTATGATCCAGATTCGAAAGAAATACCTGATGAGATGATGATGATGATGATG 720  
Db 661 CGTCAATTTATGATCCAGATTCGAAAGAAATACCTGATGAGATGATGATGATGATGATG 720  
QY 721 GTTTTGGAGAGCCTTGAACACATGACATGACATGACATGATGATGATGATGATGATG 780  
Db 721 GTTTTGGAGAGCCTTGAACACATGACATGACATGACATGATGATGATGATGATGATG 780  
QY 781 TGTGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
Db 781 TGTGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
QY 841 CAACATGAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
Db 841 CAACATGAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
QY 901 GAGGACTGGCCTTACCAAGAGATGAGCCTTGTCAAGAGAGCCTTATGATGATGATGATG 960  
Db 901 GAGGACTGGCCTTACCAAGAGATGAGCCTTGTCAAGAGAGCCTTATGATGATGATGATG 960  
QY 961 ATTCTGAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Db 961 ATTCTGAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
QY 1021 GATTAATGCTTCAACGAAGATGAGGATGAGCAAGGCTATGAGAGATCAATGCTGCAAG 1080  
Db 1021 GATTAATGCTTCAACGAAGATGAGGATGAGCAAGGCTATGAGAGATCAATGCTGCAAG 1080  
QY 1081 AATTAATGCTGCGAAGAGATGCGATGCTTTTTCAGGCCAAGCTGCCACATCAGCA 1140  
Db 1081 AATTAATGCTGCGAAGAGATGCGATGCTTTTTCAGGCCAAGCTGCCACATCAGCA 1140  
QY 1141 CCGCTAAAGAGAGAGAGCTTCCGATTAACCAAGCAGCAGCAGCAGCAGCAGCAGCAGC 1200  
Db 1141 CCGCTAAAGAGAGAGAGCTTCCGATTAACCAAGCAGCAGCAGCAGCAGCAGCAGCAGC 1200  
QY 1201 AAGGCTGCTGGAAGAGAGATTAATCTTGGATGATGATGATGATGATGATGATGATG 1260  
Db 1201 AAGGCTGCTGGAAGAGAGATTAATCTTGGATGATGATGATGATGATGATGATGATG 1260  
QY 1261 CAGGAGGAGGCTGATCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
Db 1261 CAGGAGGAGGCTGATCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
QY 1321 GGTAAAGATTTGCTCATGTTTCAAGCTGGGTAAGTGTGTTGAAGAGATGTTGGT 1380  
Db 1321 GGTAAAGATTTGCTCATGTTTCAAGCTGGGTAAGTGTGTTGAAGAGATGTTGGT 1380  
QY 1381 TACCAAGTAAACAACCTTATGCTGTTTTCCTTCTGCTTTCCTTGAAGTAAAGTAAAG 1440  
Db 1381 TACCAAGTAAACAACCTTATGCTGTTTTCCTTCTGCTTTCCTTGAAGTAAAGTAAAG 1440

QY 1441 AAAAAAAAAAAAAAAAAAAAAA 1463  
Db 1441 AAAAAAAAAAAAAAAAAAAAAA 1463

RESULT 2  
US-09-426-557-7  
Sequence 7, Application US/09426557  
Patent No. 623527  
GENERAL INFORMATION:  
APPLICANT: Mahajan, Pramod B.  
TITLE OF INVENTION: Maize Rad2/REN-1 Orthologues and Uses  
TITLE OF INVENTION: thereof  
FILE REFERENCE: 0961  
CURRENT APPLICATION NUMBER: US/09/426,557  
EARLIER FILING DATE: 1999-10-22  
EARLIER APPLICATION NUMBER: 60/112,332  
EARLIER FILING DATE: 1998-12-15  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 7  
LENGTH: 1478  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (97)...(1233)  
US-09-426-557-7

Query Match 97.2%; Score 1421.8; DB 4; Length 1478;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 7 AATAGCTCGCGCGCGGCTTCTTGGCCATCGCGGCTCAGCCGCGCCACCCGCC 66  
Db 19 AATAGCTCGCGCGCGGCTTCTTGGCCATCGCGGCTCAGCCGCGCCACCCGCC 78  
QY 67 ACAGCCGCCCGCAGAGAGATGAGGATCAAGGATGAGGATGAGGATGAGGATGAGG 126  
Db 79 ACAGCCGCCCGCAGAGAGATGAGGATCAAGGATGAGGATGAGGATGAGGATGAGG 138  
QY 127 CCCAAGGCGATGAAGAGAGAGAGATGAGGATGAGGATGAGGATGAGGATGAGG 186  
Db 139 CCCAAGGCGATGAAGAGAGAGAGATGAGGATGAGGATGAGGATGAGGATGAGG 198  
QY 187 GCCAGCATGAGCATATACAGTCTGATGATGATGATGATGATGATGATGATGATG 246  
Db 199 GCCAGCATGAGCATATACAGTCTGATGATGATGATGATGATGATGATGATGATG 258  
QY 247 ACAATATGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 306  
Db 259 ACAATATGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 318  
QY 307 TTACTGAGACCGGGAATCAAGCCAGTTTATGTTTGAATGGAAGCCCTCGATATGAG 366  
Db 319 TTACTGAGACCGGGAATCAAGCCAGTTTATGTTTGAATGGAAGCCCTCGATATGAG 378  
QY 367 AAACAGAGCTTCTGTAAGATTAAGTAAAGAGATGATGATGATGATGATGATGATG 426  
Db 379 AAACAGAGCTTCTGTAAGATTAAGTAAAGAGATGATGATGATGATGATGATGATG 438  
QY 427 GCATGAGTGAAGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 486  
Db 439 GCATGAGTGAAGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 498  
QY 487 ACAAGCAGACACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 546  
Db 499 ACAAGCAGACACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 558  
QY 547 GAGGACCTTCTGTAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 606  
Db 559 GAGGACCTTCTGTAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 618





```

DB 1141 AACAAAGAAAGAGGCTGGTGGTGAAGAGAAATATGTTGATGCTGATGATGAACCTA 1200
    |||
QY 1249 CGACTACGAACAGCAGCGGTGGCGTATCATCTTCGTTAGATTATTACTCCGTTTAA 1308
    |||
DB 1201 CGACTACGAACAGCAGCGGTGGCGTATCATCTTCGTTAGATTATTACTCCGTTTAA 1260
    |||
QY 1309 ACTCAGAGCTTGTGTAAGATTGCTCATGTTTCAAGCTGGGTAAGTTAGTTGTTG 1368
    |||
DB 1261 ACTCAGAGCTTGTGTAAGATTGCTCATGTTTCAAGCTGGGTAAGTTAGTTGTTG 1320
    |||
QY 1369 AAGAGATTGTTGATCAAGTACAAACTTATCGCTGTTT 1411
    |||
DB 1321 AAGAGATTGTTGATCAAGTACAAACTTATCGCTGTTT 1363
    |||

```

## RESULT 5

```

US-08-455-968E-2
; Sequence 2, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-455-968E-2

```

```

Query Match      20.7%  Score 302.6; DB 2; Length 1144;
Best Local Similarity 56.6%  Pred. No. 5.4e-70;
Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

```

```

QY 85 ATGGGATCAAGGCTTGTGCAAACTGCTGGCGCAATGGCCCAAGCCGATAGAGAG 144
    |||
DB 1 ATGGGATCAAGGCTTGTGCAAACTGCTGGCGCAATGGCCCAAGCCGATAGAGAG 60
    |||
QY 145 CAGAAGTTGAGAGCTACTTCCGCGCAAAATCCGCTGACGCGCAGCATGAGCATATAC 204
    |||
DB 61 AATGACATCAAGAGCTACTTGGCGTAAGTGGCCATGATGATCCTCATGAGCATTTAT 120
    |||
QY 205 CAGTCTCTGATTTGATGTAAGAGCAGCATGCAAACTTCACAAATGAGCTGGTGA 264
    |||
DB 121 CAGTCTCTGATTTGATGTAAGAGCAGCATGCAAACTTCACAAATGAGAGGAGTGA 177
    |||

```

```

QY 265 GTCACTACTCATTTGCAAGGATTTTCAACCGGACATTAAGATTACTGGAACGGGAATC 324
    |||
DB 178 ACCACACACCCAGTGAATGGATATTTCTACCGACCAATTCGATGCAAGAGGACATC 237
    |||
QY 325 AAGCAGTTTATGTTTGTGATGGCAACCTCTGATATGAAAGCAAGAGCTTGCTAA 384
    |||
DB 238 AAGCCCTGTATGTTGATGGCAACCCCGACAGCTCAAGTCAGGCGAGCTGGCCAA 297
    |||
QY 365 AGATCTCAAAAAGAGATGATGCAACCAAGATCTGACTAGGACATAGAGTAGAGAT 444
    |||
DB 298 CGCACTAGACCGCGGCTGAGGCGAGACAGACAGTGTGAGCAGGCTGTAGCTGTGGGCG 357
    |||
QY 445 AAGATGCGATGTAATAATTCAGCAAGAGAGCTGTAAAGTCAAGAGCAACAGAA 504
    |||
DB 358 GAGCGAGAGTGTGAATAATTCATAAGCGGCTGTGAAGGTCACTAAGCAGACAATGAT 417
    |||
QY 505 GATTTGTAAGCGCTTATTAAGCTTATGAGGGTTCCTGTTGATAGGACCTTGTAAGA 564
    |||
DB 418 GAGTCAAAACATCTGTGAGCTGATGGCATCCCTTATCTTATGATGACACAGTAGGCA 477
    |||
QY 565 AAGCAGATGTGAGCCCTTTGCAATAAGCAATAGGTGCTGCTTCAGAGAT 624
    |||
DB 478 GAGGCGACCTGTGCTGCGCTGTGAAGGCTGCAAAATGTATGCTGCGGCTACCGAGAC 537
    |||
QY 625 ATGACCTGCTTACTTTGAGGCGCTCAAGGCTGCTGCTGCTGCTTATTAAGTCAACTTCC 684
    |||
DB 538 ATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
    |||
QY 685 AAGAAATACCTGTGATGATGATTTGATGTTGCCAAGGTTTGGAGAGCTTGAATCACC 744
    |||
DB 598 AAAAGCTGCAATTCACCAATTCACCTGAGCGGATTCGACGAGAGCTGGGCGTGAAC 657
    |||
QY 745 ATGACAGCTTCAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 804
    |||
DB 658 CAGGAAAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
    |||
QY 805 GGTATGCGGGGCAACAGCTGTGAATCTTATGCTGACATGAGGCTCATGAAAGCATC 864
    |||
DB 718 GGTATGCGGGGCAACAGCTGTGACCTATCCAGAGCAAGGATCGAGGAGATC 777
    |||
QY 865 TTGAGAGATCTTAATTAAGACAGATATCAATTCCTAGAGAGCTGCTTCAAGAGAT 924
    |||
DB 778 GTGCGGCGACCTTGACCCCAACAGATACCTGCTGCAAAATTTGCTCCACAGAGGCT 837
    |||
QY 925 CGACCTTGTTCAGAGAGCTTAATG---TCACATTTGATATTCCTGAGCTAAATGAGCT 981
    |||
DB 838 CACGAGCTCTTCTTGAACCTGAGGTGCTGACCCAGAGCTGTGAGAGTGAAGTGGAGC 897
    |||
QY 982 GCACCTGATGAGAGAGGCTCATTAAGTTCTGCTGTAAGATATGATGTTCAAGAGAT 1041
    |||
DB 898 GAGCCAAATGAAGAGAGCTGATCAAGTTCTATGTTGTTGAAGAGAGTTCTCTGAGAG 957
    |||
QY 1042 CGGTTGCAAAAGCCATAGAGAGATCAAAATCTGCCAAGATTAATCTGTCAGAGAGAA 1101
    |||
DB 958 CGAATCCGACAGTGGGCTCAAGAGGCTGTGAAGAGCCGCAAGAGCAAGCAGAGGCGCC 1017
    |||
QY 1102 CTCGAGTCTTTTCAAGCAACTGCCACCAATCAGCACC 1142
    |||
DB 1018 CTCGATGATTTCTTCAAGCTACCGGCTCACTCTCTTCAAGC 1058
    |||

```

## RESULT 6

```

US-08-455-968E-28
; Sequence 28, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:

```

```

ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 18985-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1144 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (polynucleotide)
US-08-455-968E-28

```

```

Query Match      20.7%  Score 302.6; DB 2; Length 1144;
Best Local Similarity 56.6%; Pred No. 5,4e-70;
Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

```

```

QY 85 ATGGGATCAAGGGTTTGAGCAAACTGTCGGGACATGCGCCCAAGCGATGAAGAG 144
DB 1 ATGGGATTAAGGCGCTGGCCAAACTAATGTCGATGTGGCCCGCCATCGGAG 60
QY 145 CAGAAGTTCAGAGCTACTTCGGCGCAAAATCGCGTCGAGCGACGATGATATAC 204
DB 61 AATGACATCAAGAGCTACTTGGCCCTTAAGGTGGCCATGATGCTCTATGAGCATTTAT 120
QY 205 CAGTTCCTGATGTCAGTGAAGAGCAGCATGAATCTCAACAATGAAGCTGTGA 264
DB 121 CAGTTCCTGATGTCGTCGCCAG--GGTGGGATGTCGCAATGAGAGGGGTGAG 177
QY 265 GTCACGTGATTTGCAAGGATGTTCAACCGCAATTAAGATTAAGTCTGGAACCGGATC 324
DB 178 ACCACGACCACTGATGGGATGTTCTACCGCACCATTCGATGATGAGAGCGCATC 237
QY 325 AAGCCAGTTTATGTTTGTGCAAGCCTCTGATATGAAGAAACAGAGCTTGCTAA 384
DB 238 AAGCCGCTGATGTCCTTGTGTCAGAGCGCCACAGCTCAAGTCAGGCGCTGCGCAA 297
QY 385 AGATACCAAAAAGAGATGATGTCACCAAAAGATCTGACGAGCGTAGAGTAGAGATC 444
DB 298 CGCAGTAGCGGGGGGTGAGGCAAGAAAGACAGCTCAGACAGCTCAGCTCTGGGGCC 357
QY 445 AAAGATCGATGAAAAATTGACCAAGAGAGACTGTAAAGTGCACAGGACACACAGAA 504
DB 358 GACGAGAGAGTGAATAATTCATAAGCGCTGTGAAGTGAAGTGAAGCAGCATATAT 417
QY 505 GATGTAAAGCGCTAATTAAGCTTATGCGGGTCCCTGTGAGAGCAAGCTTCAAGCA 564
DB 418 GAGTGCAAAATGCTGAGAGCTCATAGGCGATCCCTTATCTGATGACCCAGCTAGAGA 477
QY 565 GAAGCAGAAATGTCAGCGCTTTGTCATAAAGATGAAGTGTGCTGCTTCAAGAAGAT 624
DB 478 GAGCGCAGAGTGTGCTGCTGTAAGAGGCTGCGCAAGTCTATGCTGCGGCTACGAGGAC 537
QY 625 ATGAGCTCCCTTACTTTTGGGGCTCCAGGTTCTCTTCGTCATTTAATGATCAAGTTCC 684

```

```

DB 538 ATGAGCTGCCCTCACTTGGCGAGCCCTGTGCTAATGGCAACACTGACTGGCAAGGAGCC 597
QY 685 AAGAAATATCCTGTATGAGATTTGATCTTGCCCAAGGATTTGGAGAGGCTTGAACCTACC 744
DB 598 AAAAAGCTGCGCAATCCAGGAATTCACCTGAGCGGATTTGTCAGGAGCTGGGGCTGAAC 657
QY 745 ATGAGCACTGATTAATTTGTCATCTGTCATCTGTCATCTGTCATCTGTCATCTGTCAT 804
DB 658 CAGGAACATTTTGTGATCTGTGATCTGTCATCTGTCATCTGTCATCTGTCATCTGTCAT 717
QY 805 GGTATGCGGGGCGCAACAGCTGTGAATTTTGTCAACATGSGTCCATGGAAGATC 864
DB 718 GGTATGCGGGGCGCAACAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGT 777
QY 865 TTGCAATCTTAATTAAGACAGATATCAAAATTCCTGAGACTGCGCTTACCAAGAAGCT 924
DB 778 GTGCGGCGACTTGACCCCAACAGATACCTGTGCGAGAAATTTGGCTCCACAGAGAGCT 837
QY 925 CGACGCTTGTCAAGAGAGCTAATG--TCAATTTGATTAATCTGAGCTAATATGACT 981
DB 838 CACGACCTCTTCTTGAACCTGAGGTGCTGAGACCGAGAGTCTGTGAGCTGTAAGTGAAGC 897
QY 982 GCACCTGATGAGAGGCTGCTCATTAAGTTTCCGTGAAGAAATTAATGTTCAACAGAGAT 1041
DB 898 GAGCCAAATGAAGACAGCTGATCAAGTTCAATGTTGTGTGAAGACAGTTCTCTGAGAGG 957
QY 1042 CGGTCGACAAAGGCCATGAGAGATCAATCTGCCCAAGATTAATGTCGCAAGAGAGA 1101
DB 958 CGAATCCGAGTGGGCTGCAAGAGGCTGAGTAAGACCCCAAGCAGAGACCCAGAGGCGCG 1017
QY 1102 CTCGAGTCCCTTTTCAAGCCCACTGCCACCAATCAGCACC 1142
DB 1018 CTGATGATTTCTTCAGGTGACGCGCTCAGCTCTTGAAGC 1058

RESULT 7
US-08-455-968E-9
Sequence 9, Application US/08455968E
Patent No. 5874283
GENERAL INFORMATION:
APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chih-Lin
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESSES:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 18985-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2033 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```

: TOPOLOGY: linear  
 : MOLECULE TYPE: DNA (genomic)  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 104..1237  
 : US-08-455-968E-9

Query Match 18.9% Score 276.6; DB 2; Length 2033;  
 Best Local Similarity 55.9%; Pred. No. 4.5e-63;  
 Matches 639; Conservative 0; Mismatches 479; Indels 24; Gaps 5;

QY 85 ATGGGCAATCAAGGCTTTGACAAACTGCTGGGACAAATGCGCCCAAGCGCATGAAGAG 144  
 DB 104 ATGGAAATTCACGGCCCTTGCCTTAATGCTGATGTCGCCCCAGTCCCATCGTGAAG 163  
 QY 145 CAGAAGTTCGAGCTCTTGGCGGCAAAATCGCCCGCAGCATGAGCATGATATAC 204  
 DB 164 AATGACATCAAGAGCTACTTGGTCCGCAAAAGTGGCCATGATGCTCCATGACATCTAC 223  
 QY 205 CAGTTCCTGATTTGATTTGAGAGACAGCATGAAACTCTCAAAATGAACTGGTGAA 264  
 DB 224 CAGTTCCTGATTTGATTTGAGAGACAGCATGAAACTCTCAAAATGAACTGGTGAA 280  
 QY 265 GTCACCTGATTTGACAAAGTTCACACCGGCAATAGATTTGAGAACTGGGAACTC 324  
 DB 281 ACCACAGCC---TGATGGCATGTTCTACCGCATCGC---CATGAGATGGCATC 334  
 QY 325 AAGCCAGTTATGTTTGTATGAGCAAGCCTCTGATATGAGAAACAAAGCTTCTAAA 384  
 DB 335 AAGCCTGTGACGCTTTGATGAGCAAGCCTCTGATATGAGAAACAAAGCTTCTAAA 394  
 QY 385 AGATACCTAAAGAGATGATGACCAAGATCTGACAGAGAGAGAGAGAGAGAGAT 444  
 DB 395 CGCAGTAGAT 454  
 QY 445 AAAGATGCGATTTGAAAAATTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 504  
 DB 455 GAT 514  
 QY 505 GATTTGAAAGGCTTTTAAAGACTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 564  
 DB 515 GATTTGAAAGGCTTTTAAAGACTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 574  
 QY 565 GAGCAGAT 624  
 DB 575 GAGCAGAT 634  
 QY 625 ATGAGCTCCCTTACTTTTGGGAGCTCAAGGCTTCTTCTTAAATGAGATCCAGTTCC 684  
 DB 635 ATGAGCTCCCTTACTTTTGGGAGCTCAAGGCTTCTTCTTAAATGAGATCCAGTTCC 694  
 QY 685 AAGAAATACCTGATGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGAT 744  
 DB 695 AAGAAATACCTGATGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGAT 754  
 QY 745 ATGAGCAAGTTGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGAT 804  
 DB 755 CAGAGAGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGAT 814  
 QY 805 GGTATCGGGGGGCAACAGCTGAAACTTATGCAACATGCGGCTTAAAGAGATCC 864  
 DB 815 GGTATCGGGGGGCAACAGCTGAAACTTATGCAACATGCGGCTTAAAGAGATCC 874  
 QY 865 TTGGAGATCTTAAATGAAGACAGATATCAATTCCTGAGAGAGAGAGAGAGAT 924  
 DB 875 GTGAT 934  
 QY 925 CGAGCTGTTTCAAGAT 981  
 DB 935 CAGAT 994  
 QY 962 GCACCTGATGAT 1041

DB 995 GAGCCCAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1054  
 QY 1042 CGGGTGAAGAT 1101  
 DB 1055 CGAAATTCGAT 1114  
 QY 1102 CTCGAGTCCCTTTTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1161  
 DB 1115 CTCGATGATTTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1162  
 QY 1162 TCGATTAAGAT 1221  
 DB 1163 AAGGAT 1222  
 QY 1222 T 1222  
 DB 1223 T 1223

RESULT 8  
 US-08-455-968E-4  
 ; Sequence 4, Application US/08455968E  
 ; Patent No. 5874283  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harrington, John L.  
 ; APPLICANT: Hsieh, Chih-Lin  
 ; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease  
 ; NUMBER OF SEQUENCES: 63  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, 8th Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/455,968E  
 ; FILING DATE: 30-MAY-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Apple, Randolph T.  
 ; REGISTRATION NUMBER: 36,429  
 ; REFERENCE/DOCKET NUMBER: 18985-000100  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-576-0300  
 ; TELEFAX: 415-576-0300  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1930 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: CDNA  
 ; US-08-455-968E-4

Query Match 18.6% Score 272.2; DB 2; Length 1930;  
 Best Local Similarity 55.1%; Pred. No. 6.2e-62;  
 Matches 629; Conservative 0; Mismatches 488; Indels 24; Gaps 4;

QY 85 ATGGGCAATCAAGGCTTTGACAAACTGCTGGGACAAATGCGCCCAAGCGCATGAAGAG 144  
 DB 1 ATGGGAATTCACGGCCCTTGCCTTAATGCTGATGTCGCCCCAGTCCCATCGTGAAG 60  
 QY 145 CAGAAGTTCGAGCTCTTGGCGGCAAAATCGCCCGCAGCATGAGCATGATATAC 204  
 DB 61 AATGACATCAAGAGCTACTTGGTCCGCAAAAGTGGCCATGATGCTCCATGACATCTAC 120



```

QY 205 CAGTTCCTGATTTAGTTGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CAGTTCCTGATTTAGTTGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 177
QY 265 GTCACTAGTATTTGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 ACACACAGGCTGATGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 231
QY 325 AAGCCAGTTTATGTTTATGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 AAGCCAGTTTATGTTTATGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 291
QY 385 AGATTTCTCAAAAAGATGATGCAACCAAGATGATGCAACCAAGATGATGCAACCAAGAT 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 CGCAGTGAAGAGGCGCGCGCAGGCTGAGAAAGCACTGAGAGGCTGAGAGGCTGGGATG 351
QY 445 AAGATGCGATTTAAATTTGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 GAGGAGAGGATGGAAGATGCAACCAAGATGATGCAACCAAGATGATGCAACCAAGAT 411
QY 505 GATTGTAAGGCTATTAGACTTATGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 412 GAGTGAAGACCTGCTGAGACCTCAATGGGCACTCTTACCTTGAATGACCAAGAGGCA 471
QY 565 GAAGCAAGATGTCAGACCTTTGCAATGAAGATGTCAGATGTCAGATGTCAGATGTCAGAT 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 472 GAGGCAAGCTGCTGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 531
QY 625 ATGAGCTCCCTTATTTGGGCTCCAGGCTTCTCTCAATTTAATGATTCAGATTC 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 ATGAGCTCCCTTATTTGGGCTCCAGGCTTCTCTCAATTTAATGATTCAGATTC 591
QY 685 AAGCAATTTCTGATGATGATTTGATGTCAGACGTTTGGAGAGCTTGAATCAC 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 AAGCAATTTCTGATGATGATTTGATGTCAGACGTTTGGAGAGCTTGAATCAC 651
QY 745 ATGAGCAATTTCTGATGATGATTTGATGTCAGACGTTTGGAGAGCTTGAATCAC 804
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 CAGGAGAGCTTGTGATGTCAGACGTTTGGAGAGCTTGAATCAC 711
QY 805 GATATGGGGGGGCAACACTGCAAACTTATGTCGAACATGGTCCATGAAAGATC 864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 712 GGCATTTGGCCCAAGCGGCTGTGATTCATCCAGAAACATTAAGATGAGAGATC 771
QY 865 TTGAGCAATTTCTGATGATGATTTGATGTCAGACGTTTGGAGAGCTTGAATCAC 924
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 772 GTGAGCGGCTGAGACCCAGCAAGTACCCCTTCCAGAGAACTGCTCCACAAAGAGCC 831
QY 925 CGAGCGTTTTCAGAGAGGCTTAA--TGTCACTTGTGATTTCCGAGCTTAAATGACT 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 832 CAGTACGCTTCTGAGAGGCTGAGAGTGTGACCCAGAGCTTGTGAGAGCTGAGAGC 891
QY 982 GCACCTGATGAGAGGCTGCTCAATAGTTTCTGTTAAAGATTAATGTTTCAAGAGAT 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 892 GAGCAATTTAAAGAGATGTTGTCATTAATTTATGTTGTTAAAGAGATTTTCTGAAG 951
QY 1042 CGGATGCAAGAGGCTTAAAGATGATGATGTCAGAGATTAATGTTTCAAGAGAT 1101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 952 CGAATTTGCGAGTGGGCTCAAGGCTGAGTATGAGAGCCCGAGGAGAGACCCAGAGC 1011
QY 1102 CTGAGTCTTTTTCAGGCAACTGTCACACATGAGCAGCGCTTAAAGAGAGACT 1161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1012 CTGAGTCTTTTTCAGGCAACTGTCACACATGAGCAGCGCTTAAAGAGAGACT 1059
QY 1162 TCGATTTAAAGCAAGAGAGCTGCAAGAAACAAAGAGCTGTTGAAAGAGAA 1221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1060 AAGGAGCAACACCAAGAGGCTGCTAAGAGAAAGCAAAAGACTGAGGAGCGGAG 1119
QY 1222 T 1222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1120 T 1120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 9
US-08-455-968E-6
; Sequence 6, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-May-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ. ID NO. 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-455-968E-6

Query Match      18.5%; Score 270.2; DB 2; Length 1149;
Best Local Similarity 55.5%; Pred. No. 1,7e-61;
Matches 599; Conservative 0; Mismatches 453; Indels 27; Gaps 3;

QY 85 ATGGGCAATCAAGGATTTGACGAACCTGCGGACAAATGCGCCGACAGCCAGATGAGATATAC 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGGGCAATCAAGGATTTGACGAACCTGCGGACAAATGCGCCGACAGCCAGATGAGATATAC 60
QY 145 CAGAGTTCGAGAGCTACTTGGCGGCAAAATCGCCGCGACAGCCAGATGAGATATAC 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AGCGATTTCAAGAGCTTTTGGCAGAAAGTTGCGATGAGGCTCTATCTCTATAT 120
QY 205 CAGTTCCTGATTTAGTTGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CAGTTCCTGATTTAGTTGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 180
QY 265 GTCACTAGTATTTGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 ACACGTCACACTGATGAGGATGTTTATGAGACACTGAGATGATGATGATGATGATGATGAT 240
QY 325 AAGCCAGTTTATGTTTATGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 AAGCCAGTTTATGTTTATGGAAGACAGGACATGGAACCTCTCAACAATGAACTGGTGA 300
QY 385 AGATTTCTCAAAAAGATGATGCAACCAAGATGATGCAACCAAGATGATGCAACCAAGAT 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 CGTCTTCAAGAGGCTGCAAGAAACAAAGATGCAAGAGGCTGCAAGAGGCTGCAAGAGG 360
QY 445 AAGATGCGATTTGAAATTTGAGCAAGAGGCTGCAAGAGGCTGCAAGAGGCTGCAAGAGG 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 AAGATGCAAGCAAGAAAGG-----AGATTTGGAAGGCTCAAAAGAGCATATGAA 411

```



```

QY 505 GATTGTAACGGCTATTAGACTTATGGGGTTCCTGTTGAGAGCACTTCTGANGCA 564
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 412 GAAGCCCAAAAATTACTAGACTAATGGAAATCCATTAATAAGAGCCCAACGAAAGCT 471
QY 565 GAAGCAGAAATGCGAGCCCTTGCATAAACGATAGAGTTCGCTGTTGCTTCAGAGAT 624
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 472 GAGCTCAATGCTGAGTGTGCAAGAAAGGAAAGGTGTATGCGGACGAGAGTGAAGAT 531
QY 625 ATGAGTCCCTTACTTGTGGGCTCCAGGGTTCCTGCTCATTTAATGATCCAGTTCC 684
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 532 ATGACACACTGTATTGAACACCCCTTCTGTTGAACATTTGACTTTTCAGAGGCC 591
QY 685 AAGAAATACCTGTGATGAAATTTGATGTCGCAAGTTTGGAGAGCTTGAAGCACC 744
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 592 AAGAAGAACCCGATTACGAATAATAGTACTGAATAGTTTGAAGAGACTGACTTGACA 651
QY 745 ATGAGACCATTTCTATTGATTTTGTGATCCCTGTGTGATGTGACTATTTGATAGCATCAA 804
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 652 ATGAGACAGTTTGTGATCTTTGCAATATGCTGTTGTGTGACTGCTGGAAGCATAGA 711
QY 805 GGTATCGGGGGGCAACAGCTCTGAACTTATTCGTCAACATGGGTCCATGGAAGCATC 864
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 712 GGTGTTGTCACAGTACAGCCCTTAATATGATAAAGACATGATCCATGAGAAATATC 771
QY 865 TTGGAGAACTT-----ATAAGACATATCAAAATTCCTGAGAGACTGG 909
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 772 GTGAGATTATTGAATCTGGGAGTCACAACTAATATGGAATCCCGAAGAGCTGG 831
QY 910 CCTTACCAAGAAAGCTGACCTGTTGTCAGAGACCTTAATGTCACATTGGATA--TTCT 966
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 832 CCTTACCAAGAAAGCATGCTGTTTCTGACCTGAAGTTATGATGATGAGCAATA 891
QY 967 GAGCTAAATGAGCTGACCTGATGAGAGGCTGTCATAGTTTCCCTGTAAGATTAAT 1026
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 892 AACTGAAATGTCGCCACCAAGAGAAAGAACTTATGAGATTTATGATGATGAG 951
QY 1027 GATTTCACAGACATCGGTGACAAAGGCCATAGAGAAATCAATCTGCCAAGATPAA 1086
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 952 AAATTCAGTGAAGAAAGAGTAACTGTATATCAATGAAATGAAAGAGCTTGAATCT 1011
QY 1087 TCGTCGCAAGAGAGTCTGAGTCTTTTCAAGCCAGCTGCGACCATGACCCGCT 1145
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1012 GGCATTTCAGGATGATTAGATGGGTTCTTCCAGATGGTGCCTAAGACCAAGAACAGCT 1070

RESULT 10
US-08-757-653-175
; Sequence 175, Application US/08757653
; Patent No. 5843569
; GENERAL INFORMATION:
; APPLICANT: Kaiser, Michael W.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Lyamichev, Natasha
; TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
; NUMBER OF SEQUENCES: 190
; TITLE OF INVENTION: The most stable FEN-1 Endonucleases
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,653
; FILING DATE:
; CLASSIFICATION: 435

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02565
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 175:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1020
; US-08-757-653-175

Query Match 10.9%; Score 159.6; DB 2; Length 1023;
Best Local similarity 54.6%; Pred. No. 1.5e-32;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 134 CATGACAGGAGAGAGTTGAGAGTACTGCGCGCAAAATCGCGTCAGCCAGCA 193
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 29 CAAGAAAGAAATTGATGTAGAAAACCTATACGGGAAAAAATCGCATGCGCTCTTA 88
QY 194 TGAGCATATACCATGTTCTGTATGTTGAGAGACAGCATGAAACTCTCAAAATG 253
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 89 ATGCATCTACCAATTTTGTCCAAATTAAGACAGAAAGATGAACTCCACTGTGAT 148
QY 254 AAGCTGTGAAGTCTCATGATTTGCAAGAAATGTTCAACCGACATTAAGATTAAGT 313
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 149 CAAGAGTATGATATACCTCCACCTAAGCGGCTCTTTACAGACATTAACCTAATG 208
QY 314 AAGCGGAATCAAGCACTTTATTTTGTATGCAAGCCTCTGATATGAGAAACAAG 373
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 209 AAGCTGGAATTAACCTGTGATTTTGTATGAGAACTCCAGATTCATAAAGAAAG 268
QY 374 ACCCTGTAAGATACCTCAAAAGAGTGAATGACAAACAGATCTGAGTACAGTAG 433
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 269 AACTGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 328
QY 434 AGGTAGAGATTAAGATCGATGTAAGAAATAGACCAAGAGACTGTAAAGTACAAAGC 493
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 329 AAAAAGGAGAGATGAGAGAGCAAGAAATATGCCCAGAAAGACCAAGGTAATGAAA 388
QY 494 AACACACGAAAGATTGTAACGGCTATTAAAGCTTATGAGGGGTTCCTGTTAGAGGCAC 553
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 389 TCGTATGAGAGATGCAAAAACCTTAGAGCTTATGAGGATTCCTATGTTCAAGCAC 448
QY 554 CTTTGAAGCAGACCAACAAATGTGAGCCCTTTTGCATTAAGCATAGGTTTCCGTGTTG 613
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 449 CTAGCGAGGAGAGGCCCAACGTCATATATGCGCGGAAAGGAGGAGGCTATGATCGG 508
QY 614 CTTGCAAGATATGAGCTCCCTTACTTTTGGGCTCCACAGGTTCTCTGTCATTTATG 673
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 509 CTAGTCACATATGATTCCTTACTTTTGGAGCTCCAGAGCTGTAGAAACTTAACA 568
QY 674 ATCCAAATTCCAAGAAATACCTGTGATGATGATTTGATGTTG 715
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 569 TAACAGAGAAAAAGAAAGTTGCTGGAAAAAATGTCTACGTGG 610

RESULT 11
US-08-823-516-78
; Sequence 78, Application US/08823516
; Patent No. 5984069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Mary Ann D.
; APPLICANT: Brow, Mary Ann D.

```



SEQUENCE CHARACTERISTICS:  
LENGTH: 1023 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1020  
US-08-759-038-114

Query Match 10.9%; Score 159.6; DB 3; Length 1023;  
Best Local Similarity 54.6%; Pred. No. 1.5e-32;  
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 134 CGATGAAGAGCAGACGATGAGAGCTACTTGGCGCGCAAAATCGCGCTGACGCCAGCA 193  
DB 29 CAGAAAAGAAATGAGTTAGAAAACCTATACGGAAAAAATCGCAATCGAGCTCTTA 88  
QY 194 TGACATATACCACTTCTGATTGTAGTTGGAAGACAGCATGAACTCTACAAATG 253  
DB 89 ATGCATCTACCAATTTTTCACAAATAGACAGAAAGTGAACCTCCACTATGATG 148  
QY 254 AAGCTGGAAGTCACTAGTCACTTTGCAAGCAATGTTCAACGCAATATAGTACTG 313  
DB 149 CAAAGGTAAGATTAACCTCCACTAGCGGGCTCTTTTACAGACAAATTAACCTAATG 208  
QY 314 AAGCGGAATCAAGCACTTATGTTTATGCAAGCTCTCTGATATGAAGAACAG 373  
DB 209 AGGCTGGAATTAACCTGATGTTTATGTTTATGAGAACTCCAGATTCAAAAAGAA 268  
QY 374 AGCTTGTAAAGATTAAGTCAAAAGAGATGATCAACCAAGATCTGATGAGGAGT 433  
DB 269 AGCTGAAAAAGAAAG 328  
QY 434 AGGTGAGGATTAAGATGACATTAATAATGAGCAAGAGAGAGAGAGAGAGAGAG 493  
DB 329 AAAAAAGAGATTAAG 388  
QY 494 AACACAGAGATTAAG 553  
DB 389 TGCTCATGAGATGACAAAAAACTTTAGAGCTTATGGAATTCCTTAATGTTCAAGC 448  
QY 554 CTTTGAAGCAGAGCAAGATGTCACCTTTGCAATTAAGATTAAGAGAGAGAGAG 613  
DB 449 CTAGGAG 508  
QY 614 CTTCAAGAGATGAGTCTCTTACTTTTGGGCTCCAGCGTTCTTCTGATTAATG 673  
DB 509 CTAGTCAAGATTAAGATGAGTCTCTTACTTTTGGAGCTCCAGAGCTTTTGAAC 568  
QY 674 ATCCAGTTCCAGAAAAATACCTGATGATGATTAATGATGTTG 715  
DB 569 TACAGAGAAAAAGAAAGTTGCTGAGAAAAATGCTACGTCG 610

RESULT 13  
US-08-758-314-114  
Sequence 114, Application US/08758314  
Patent No. 6090606  
GENERAL INFORMATION:  
APPLICANT: Kaiser, Michael W.  
APPLICANT: Lymanchev, Victor I.  
APPLICANT: Lymanchev, Natasha  
TITLE OF INVENTION: Improved Cleavage Agents  
NUMBER OF SEQUENCES: 134  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Medlen & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States Of America  
ZIP: 94104

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/758,314  
FILING DATE: 02-DEC-1996  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/  
FILING DATE: 29-NOV-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/682,853  
FILING DATE: 12-JUL-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/599,491  
FILING DATE: 24-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Ingolia, Diane E.  
REGISTRATION NUMBER: 40,027  
REFERENCE/DOCKET NUMBER: FORS-02575  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ. ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1023 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1020  
US-08-758-314-114

Query Match 10.9%; Score 159.6; DB 3; Length 1023;  
Best Local Similarity 54.6%; Pred. No. 1.5e-32;  
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 134 CGATGAAGAGCAGAGATGAGAGCTACTTGGCGCGCAAAATCGCGCTGACGCCAGCA 193  
DB 29 CAGAAAAGAAATGAGTTAGAAAACCTATACGGAAAAAATCGCAATCGAGCTCTTA 88  
QY 194 TGACATATACCACTTCTGATTGTAGTTGGAAGACAGCATGAACTCTACAAATG 253  
DB 89 ATGCATCTACCAATTTTTCACAAATAGACAGAAAGTGAACCTCCACTATGATG 148  
QY 254 AAGCTGGAAGTCACTAGTCACTTTGCAAGCAATGTTCAACGCAATATAGTACTG 313  
DB 149 CAAAGGTAAGATTAACCTCCACTAGCGGGCTCTTTTACAGACAAATTAACCTAATG 208  
QY 314 AAGCGGAATCAAGCACTTATGTTTATGCAAGCTCTCTGATATGAAGAACAG 373  
DB 209 AGGCTGGAATTAACCTGATGTTTATGTTTATGAGAACTCCAGATTCAAAAAGAA 268  
QY 374 AGCTTGTAAAGATTAAGTCAAAAGAGATGATCAACCAAGATCTGATGAGGAGT 433  
DB 269 AGCTGAAAAAGAAAG 328  
QY 434 AGGTGAGGATTAAGATGACATTAATAATGAGCAAGAGAGAGAGAGAGAGAGAG 493  
DB 329 AAAAAAGAGATTAAG 388  
QY 494 AACACAGAGATTAAG 553  
DB 389 TGCTCATGAGATGACAAAAAACTTTAGAGCTTATGGAATTCCTTAATGTTCAAGC 448  
QY 554 CTTTGAAGCAGAGCAAGATGTCACCTTTGCAATTAAGATTAAGAGAGAGAGAG 613  
DB 449 CTAGGAG 508



```

Qy 614 CTTGAGAGATATGACACCCCTACTTTGGGGCTCCACGGTCCCTGCTCATTTAATGG 673
Db 509 CGAGTCAGATTATGATTCACACTCTTGGTGCTCCAGGTGATTAGGAATCTGACAA 568
Qy 674 ATCCAAGTTCACAGAAAATACCTGTGATGGAATTTGATGTTG 715
Db 569 TTACGGGAAAAGAAAGATGCCCTGGGAAAGATGTTACGTTG 610

Search completed: November 5, 2002, 11:02:54
Job time : 46.2144 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 10:28:52 : Search time 1251.78 seconds  
(without alignments)  
15774.429 Million cell updates/sec

Title: US-09-805-311-1  
Perfect score: 1463  
Sequence: 1 caccagatagctcgcgcgc.....aaaaaaaaaaaaaaaa 1463

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_esthu:\*  
4: em\_estnu:\*  
5: em\_estcov:\*  
6: em\_estcpl:\*  
7: em\_estcro:\*  
8: em\_estl:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778.8	53.2	901	10	BG837708 Zm10.01f08
2	560.8	38.3	586	9	AI881599 606068G09
3	536.4	36.7	550	10	BE639422 946033A02
4	534.4	36.5	534	10	BE639421 946033A02
5	517.8	35.4	533	9	AM562789 660065H06
6	501	34.2	532	9	AM000375 614014D03
7	465	31.8	648	9	AV913663 AV913663
8	463.4	31.7	474	9	AM559173 660065H06
9	462	31.6	470	9	AI861468 614014D03
10	448	30.6	467	9	BE186786 946012C08
11	429.8	29.4	475	9	AM562517 660065H06
12	420.4	28.7	733	9	AM774700 EST333851
13	413	28.2	553	9	AI834484 606068G09
14	411.6	28.1	539	10	BG263233 WHE2339_F
15	402.2	27.5	414	9	AM288784 707010C02
16	398	27.2	456	9	AI065689 ag91f12.x
17	378.8	25.9	632	9	AM585913 EST317536

18	372.6	25.5	609	10	BG096545	BG096545	EST7461064
19	370.4	25.3	754	10	BE416584	BE416584	MUG0009.BI
20	369.6	25.3	820	10	BG414505	BG414505	HVSMEX000
21	364.6	24.9	645	9	AM691681	AM691681	NFO43B09S
22	348	23.8	460	10	BE417817	BE417817	MUG024.H1
23	347.2	23.7	683	9	AV918966	AV918966	AV918966
24	346.8	23.7	454	9	AV915644	AV915644	AV915644
25	341.8	23.4	363	9	AM562518	AM562518	660065H06
26	325.4	22.2	539	10	BF098320	BF098320	EST428761
27	317.2	21.7	470	9	AM288831	AM288831	707010E11
28	314.8	21.5	376	9	AI065346	AI065346	ag88602.x
29	314.8	21.5	579	9	AV833393	AV833393	AV833393
30	314.4	21.5	592	9	AI780966	AI780966	EST261845
31	308	21.1	526	10	BI419720	BI419720	LJNEST479
32	300.6	20.5	695	9	BE052579	BE052579	GA_Ea000
33	273	18.7	549	9	AM696665	AM696665	NE109F01S
34	272.4	18.6	942	9	AL560007	AL560007	AL560007
35	266.2	18.1	903	9	AL531350	AL531350	AL531350
36	262.4	17.9	851	9	AL519300	AL519300	AL519300
37	248.6	17.0	689	10	BE311755	BE311755	601143566
38	245.4	16.8	769	10	BG574950	BG574950	602598038
39	244.4	16.7	660	10	BJ011893	BJ011893	AL560395
40	242.8	16.6	806	9	AL560395	AL560395	AL560395
41	237.6	16.2	762	10	BG337603	BG337603	602435150
42	234.6	16.0	866	10	BI827898	BI827898	603073821
43	233.6	16.0	704	9	AM783795	AM783795	rp09907.y
44	233.4	16.0	963	12	CNS06WXR	AL419029	T3 end of
45	233	15.9	785	10	BG756459	BG756459	602715791

## ALIGNMENTS

RESULT 1  
LOCUS BG837708 901 bp mRNA linear EST 25-MAY-2001  
DEFINITION Zm10.01f08\_A Zm10.AAFC-ECORC\_Fusarium\_graminearum\_corn\_silk zea  
mays CDNA clone Zm10.01f08, mRNA sequence.

ACCESSION BG837708  
VERSION BG837708.1 GI:44204031  
KEYWORDS EST.

SOURCE Zea mays.  
ORGANISM Zea mays.

REFERENCE Harris,L.J., Balcerzak,M., Allard,S., Sapiro,A., Couroux,P., De Moors,A., Hattori,I., Ouellet,T., Robert,L.S., Singh,J.A., Sprott D., and Finkler,N.A.  
Expressed Sequence Tags from Maize Silk Six Hours After Silk Channel Inoculation with Fusarium graminearum

JOURNAL Contact: Harris, Linda J.  
Unpublished (2001)  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-Food Canada  
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
CANADA  
Tel: (613) 759-1314  
Fax: (613) 759-6566  
Email: harris@em.agr.ca.

## FEATURES

location/Qualifiers  
1..901  
/organism="Zea mays"  
/cultivar="CO388"  
/db\_xref="taxon:457"  
/clone="Zm10.01f08"  
/clone\_id="Zm10.AAFC-ECORC\_Fusarium\_graminearum\_corn\_silk"  
/tissue\_type="Silk"  
/dev\_stage="4-5 days post-silk emergence"  
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI; Site 2: XhoI; field-grown corn was silk channel-inoculated"





Db	435	CTGAGGAGTAGAGTAGAGATTAAGATGCGATTGAAAAATTGAGCAAGAGCATTGTA	494
Oy	482	AGGTCACAAAGGCACACACAAAGATTGTAACGGCTATTAGACTTATGGGGGTTCTGG	541
Db	495	AGGTCACAAAGGCACACACAAAGATTGTAACGACATTAAGACTTATGGGGGTTCTGG	554
Oy	542	TTGTAGAGGACACTTCTGAAGCAGAGAGAA	573
Db	555	TTGTAGAGGACACTTCTGAAGCAGAGAGAGAA	586

RESULT	3	
LOCUS	BE639422	
DEFINITION	BE639422	550 bp mRNA
ACCESSION	U46033A02.v2	linear EST 30-AUG-2000
VERSION	BE639422	tassel primordium prepared by Schmidt lab Zea
KEYWORDS	BE639422.1	GI:9952839
SOURCE	EST.	
ORGANISM	Zea mays.	
	Zea mays.	
	Zea mays.	

REFERENCE	1
AUTHORS	Walbot, V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL COMMENT

Unpublished (1999)  
Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 946033 row: A column: 02.

FEATURES	Location/Qualifiers
source	1. .550

```

/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassal primordium prepared by Schmidt
lab"
/issue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI;
Site 2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybriZAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."

```

	Query Match	36.7%	Score 536.4	DB 10	Length 550
	Best Local Similarity	99.6%	Pred. No. 1.6e-66		
	Matches 548	Conservative 0	Mismatches 1	Indels 1	Gaps 1
QY	46 AGCGCGCCGCCGCGACGCCCGCAGCGAGCGATGGCGATCAAGGGTTTGACG	105			
Db	1 AGCGCGCGCGCGCGACCGCGCANCGCGCGGAGCGAGATGGCGATCAAGGGTTTGACG	60			
QY	106 AAACGTGCGCGACATGCGCCCAAGGCGATGAAGCGAGCGAAGTTGCAAGCTACTTC	165			
Db	61 AAATGCTGGCGGACATGCGCCCAAGGCGATGAAGGCGAGAGTTGCAAGCTACTTC	120			
QY	166 GCGCGCAAAATGCGGTGAGCGCGCAAGCATATCCAGTCCGATTTGAATTTGGA	225			
Db	121 GCGCGCAAAATGCGGTGAGCGCGCAAGCATATCCAGTCCGATTTGAATTTGGA	180			

OY	226	AGGACAG -CATGAAACTCTCACAAATGAGCTGGGAAAGTCACTAGCATTTGCAAG	284
Db	181	AGAGCTGTGTAATGGAACCTCACAAATGAAAGCTGTGAAGTCACTAGCATTTGCAAG	240
OY	285	AATGTTCAACCGGACATATAGATTCTGTAAGCGGGAGATCAAGCCAGTTATGTTTTGA	344
Db	241	AATGTTCAACCGGACATATAGATTCTGTAAGCGGGAGATCAAGCCAGTTATGTTTTGA	300
OY	345	TGGCAACCTCTCGATATATGAAGCAAGACCTTGCTAAAGATACTCCAAAAGAGATGA	404
Db	301	TGGCAACCTCTCGATATATGAAGCAAGACCTTGCTAAAGATACTCCAAAAGAGATGA	360
OY	405	TGCAACCAAGATCTGACTGAGGCGAGTAGGTGAGATTAAGATGCGATTGAAAAATT	464
Db	361	TGCAACCAAGATCTGACTGAGGCGAGTAGGTGAGATTAAGATGCGATTGAAAAATT	420
OY	465	GAGCAAGAGACTGTAAAGTCAACAAGGCACACACGAAGATTTGAACGCTATTAAAG	524
Db	421	GAGCAAGAGACTGTAAAGTCAACAAGGCACACACGAAGATTTGAACGCTATTAAAG	480
OY	525	ACTTATGGGGTTCCTGTTGTAGAGGACCTTTGGAAGCAAGACAGAATGTCAACCCCT	584
Db	481	ACTTATGGGGTTCCTGTTGTAGAGGACCTTTGGAAGCAAGACAGAATGTCAACCCCT	540
OY	585	TTTGATTAAC	594
Db	541	TTTGATTAAC	550

RESULT 4	BE639421/c	BE639421	554 bp	mrna	linear	EST 30-AUG-2000
LOCUS	BE639421/c					
DEFINITION	946003JA02.x2 946 - tasseli					
	primordium prepared by Schmidt lab					
	Zeae					
ACCESSION	BE639421					
VERSION	BE639421.1					
KEYWORDS	EST.					
SOURCE	Zeae					
ORGANISM	Zeae					

REFERENCE	1 (bases 1 to 554)
AUTHORS	Walbot, V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL	University
COMMENT	Unpublished (1999)
	Contact: Walbot V

Department of Biological Sciences  
Stanford University  
855 California Ave,  
Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walp0t@stanford.edu  
Plate: 946033 row: A column: 02.

```

FEATURES
source
Location/Qualifiers
1..554
/organism="zebra mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassels primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/notes="Organ: tassels; Vector: HybriZAP; Site:1; Ecoli:
site:2; Host: George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybriZAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."
BASE COUNT
145 a 131 c 110 g 168 t
ORIGIN

```

```

Query Match      36.5%; Score 534.4; DB 10; Length 554;
Best Local Similarity 99.8%; Pred. No. 3e-68;
Matches 535; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 896 TTCCGAGAGCTGGCGCTTCCCAAGAGCTCGAGCGCTTGTTCAGAGAGCGCTTAATGTCACAT 955
    |||||||
Db 554 TTCCGAGAGCTGGCGCTTCCCAAGAGCTCGAGCGCTTGTTCAGAGAGCGCTTAATGTCACAT 495

OY 956 TGGATATTCCTGAGCTAAATATGACTGCACCTGATGAGAGAGGCTCTCAATGTTCTCTGG 1015
    |||||||
Db 494 TGGATATTCCTGAGCTAAATATGAGGCTGCACCTGATGAGAGAGGCTCTCAATGTTCTCTGG 435

OY 1016 TAAAGATATATGTTTCAACGAGAGATCGGGTACAAAGAGCGCTTAGAGAGATCAATCTG 1075
    |||||||
Db 434 TAAAGATATATGTTTCAACGAGAGATCGGGTACAAAGAGCGCTTAGAGAGATCAATCTG 375

OY 1076 CCAGAGATTAATGCTGCAAGAGAGACTGAGTCTTTTCAAGCCAACTGCCACACAT 1135
    |||||||
Db 374 CCAGAGATTAATGCTGCAAGAGAGACTGAGTCTTTTCAAGCCAACTGCCACACAT 315

OY 1136 CACGACCGCTAAAGAGAGAGACTTGGATTAATAACAAGCAGCAGCTCGAACAAGA 1195
    |||||||
Db 314 CACGACCGCTAAAGAGAGAGACTTGGATTAATAACAAGCAGCAGCTCGAACAAGA 255

OY 1196 AAACAAAGGCTGTGGAAGAGAAATTAATCTGATGCTTGATGACTAGACTAGACTAC 1255
    |||||||
Db 254 AAACAAAGGCTGTGGAAGAGAAATTAATCTGATGCTTGATGACTAGACTAGACTAC 195

OY 1256 GAAGAGAGCGGTGGCTGATCACTTGGTATTAATTAATTAATTAATTAATTAATTAATTA 1315
    |||||||
Db 194 GAAGAGAGCGGTGGCTGATCACTTGGTATTAATTAATTAATTAATTAATTAATTAATTA 135

OY 1316 GCTTGTGTAAGTGTGCTGATGTTTCAAGCTGGGGTAAGTTAGTTGTTTGAAGAGAT 1375
    |||||||
Db 134 GCTTGTGTAAGTGTGCTGATGTTTCAAGCTGGGGTAAGTTAGTTGTTTGAAGAGAT 75

OY 1376 TGGGTACCAAGTAACAAACTATGCGTGTGTTTACTCTTGTCTTGAAGTA 1431
    |||||||
Db 74 TGGGTACCAAGTAACAAACTATGCGTGTGTTTACTCTTGTCTTGTGAAGTA 19

RESULT 5
AM562789 553 bp mRNA linear EST 10-MAR-2000
LOCUS 660065H06.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AM562789
VERSION AM562789.1 GI:7216667
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays;
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 553)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY Unpublished (1999)
CONTACT: Walbot V
DEPARTMENT: Department of Biological Sciences
STANFORD UNIVERSITY
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.
FEATURES
source 1..553
location/Qualifiers
1..553 /organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"

```

```

/clone_lib="660 - Mixed stages of anther and pollen"
/rissue_type="whole premiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="X10LR"
/note="Organ: anthers; Vector: Lambda Zap; Site: 1; EcoRI;
Site: 2; XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 171 a 111 c 129 g 142 t
ORIGIN

Query Match      35.4%; Score 517.8; DB 9; Length 553;
Best Local Similarity 98.7%; Pred. No. 7.6e-66;
Matches 522; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 903 GGACGCGCTTACCAAGAGAGCTCGACGCTTGTTCAGAGAGCGCTTAATGTCACATTGGATAT 962
    |||||||
Db 4 GGACGCGCTTACCAAGAGAGCTCGACGCTTGTTCAGAGAGCGCTTAATGTCACATTGGATAT 63

OY 963 TCCTGAGCTAAATATGAGTGCACCTGATGAGAGAGGTCCTCAATGTTCTCTGTAAGA 1022
    |||||||
Db 64 TCCTGAGCTAAATATGAGTGCACCTGATGAGAGAGGTCCTCAATGTTCTCTGTAAGA 123

OY 1023 TATGTTTCAACGAGAGATCGGGTGACAAGCCATAGAGAGATCAATCTGCCAAGAA 1082
    |||||||
Db 124 TATGTTTCAATGAGAGATCGGGTGACAAGCCATAGAGAGATCAATCTGCCAAGAA 183

OY 1083 TAAATGCTGCAAGAGAGAGCTGAGTCTTTTCAAGCCAACTGCCACCAATCAGACAC 1142
    |||||||
Db 184 TAAATGCTGCAAGAGAGAGCTGAGTCTTTTCAAGCCAACTGCCACCAATCAGACAC 243

OY 1143 GCTAAACGAGAGAGACTTGGATTAATAACAAGCAGCAGCTCGAACAAGAGAAACAA 1202
    |||||||
Db 244 GCTAAACGAGAGAGACTTGGATTAATAACAAGCAGCAGCTCGAACAAGAGAAACAA 303

OY 1203 GGTGTGGAAGAGAGAAATATATCTTGATGCTTGATGATGATGATGATGATGATGATGAT 1262
    |||||||
Db 304 GGTGTGGAAGAGAGAAATATATCTTGATGCTTGATGATGATGATGATGATGATGATGAT 363

OY 1263 GCGGTGCGGTGATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1322
    |||||||
Db 364 GCGGTGCGGTGATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 423

OY 1323 TAAAGTTTGCATGTTTCAAGCTGGGGTAAGTTAGTTGTTTGAAGAGATTGGTGA 1382
    |||||||
Db 424 TGAAGTTTGCATGTTTCAAGCTGGGGTAAGTTAGTTGTTTGAAGAGATTGGTGA 483

OY 1383 CCAGTAACAAACTATGCGTGTGTTTACTCTTGTCTTGAAGTA 1431
    |||||||
Db 484 CCAGTAACAAACTATGCGTGTGTTTACTCTTGTCTTGTGAAGTA 532

RESULT 6
AM000375 532 bp mRNA linear EST 08-SEP-1999
LOCUS 614014D03.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AM000375
VERSION AM000375.1 GI:5847296
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays;
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 532)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY Unpublished (1999)
CONTACT: Walbot V
DEPARTMENT: Department of Biological Sciences
STANFORD UNIVERSITY

```



Df	114	TGCGCTTAGATTTAATTAACTCCCTGTTTTAAGTCAACACTTGCTGGAAAGTTTCGCATCT	55
OY	1340	TTTCAACTGGGTGAAGTAACTAGTGTCTTGGAAGACAATGGTGTACCAAGTACCAA	1393
Df	54	TTCAACTGGGTGAAGTAACTAGTGTCTTGGAAGACAATGGTGTACCAAGTACCAA	1

  

RESULT 9	A1861468	470 bp	mRNA	linear	EST 19-JUL-1999
LOCUS	A1861468/c	614014D03.xl 614	- root cDNA library from Walbot Lab	Zea mays cDNA.	
DEFINITION	mRNA sequence.				
ACCESSION	A1861468				
VERSION	A1861468.1	GI:5525575			
KEYWORDS	EST.				
SOURCE	Zea mays.				
ORGANISM	Zea mays.				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC				
TITLE	clade: Panicoideae; Andropogoneae; Zea.				
JOURNAL	1 (bases 1 to 470)				
COMMENT	Walbot V.				
	Maise ESTs from various cDNA libraries sequenced at Stanford University				
	Unpublished (1999)				
	Contact: Walbot V				
	Department of Biological Sciences				
	Stanford University				
	855 California Ave, Palo Alto, CA 94304, USA				
	Tel: 650 723 2227				
	Fax: 650 725 8221				
	Email: walbot@stanford.edu				
	Plate: 614014 row: D column: 03.				
FEATURES	Location/Qualifiers				
source	1..470				
	/organism="Zea mays"				
	/cultivar="W23"				
	/db_xref="taxon:4577"				
	/clone_lib="614"-root cDNA library from Walbot Lab"				
	/tissue_type="root"				
	/tissue="Lysate"				
	/dev_stage="3-4 days old"				
	/lab_host="X10LR"				
	/note="Organ: root; Vector: pBluescriptII SK+; Site_1:				
	ECORI; Site_2: XhoI; 3-4 days old root tissue from Walbot				
	Lab (LM)"				
BASE COUNT	112 a	113 c	93 g	152 t	
ORIGIN					
Query Match	31.6%	Score 462;	DB 9;	Length 470;	
Best Local Similarity	98.9%;	Pred. No. 9, 5e-56;			
Matches 465;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;	

  

OY	937	AAGGACCTAATGTCACATGATATTCCTGAGCTAAATGAGTGCACCTGATAGGAG	996
Df	470	AAGGAGACTAATGTCCCATGATGATATTCCTGAGCTAAATGAGTGCACCTGATAGGAG	411
OY	997	GGTCTATAGTATTCCTGTTAAATAATATGTTTCAACGAAAGATCGGTAAGAAGGCC	1056
Df	410	GGTCTATAGTATTCCTGTTAAATAATATGTTTCAACGAAAGATCGGTAAGAAGGCC	351
OY	1057	ATGAGAGAAATCAAAATCTGCCAAGAAATTAATGTCGCAAGGAAGATCGGTCCTTTTTC	1116
Df	350	ATGAGAGAAATCAAAATCTGCCAAGAAATTAATGTCGCAAGGAAGATCGGTCCTTTTTC	291
OY	1117	AAGCCAATGCGCACACATCAGACCGCTAAACGGAAGAGACTTGGATPAAAACAAGC	1176
Df	290	AAGCCAATGCGCCCCACATCAGACCGCTAAACGGAAGAGACTTGGATPAAAACAAGC	231
OY	1177	AAGGACGCTGCGCAACGAACAAAGGCTGTTGGAAGAGAAATATATCTTGATGCTT	1236
Df	230	AAGGACGCTGCGCAACGAACAAAGGCTGTTGGAAGAGAAATATATCTTGATGCTT	171

QY 1237 GATGACAACTACGACAGCAAGAGCGGTGGCTGATCATCTTGGTTAGATTATTTAA 1296  
 |||||||  
 Db 170 GATGTCAAACTACGACAGCAAGAGCGGTGGCTGATCATCTTGGTTAGATTATTTAA 111  
 |||||||  
 QY 1297 CTCCCTGTTTACTACGAGCTTTGTTAAAGTTTGCTCATGTTTCAAGCGGGTAACT 1356  
 |||||||  
 Db 110 CTCCCTGTTTAACTACGAGCTTTGTTAAAGTTTGCTCATGTTTCAAGCGGGTAACT 51  
 |||||||  
 QY 1357 TACTGTGTTTAAAGAGATGTTGTACCAAGTAACTAATCTATCGCTGT 1406  
 |||||||  
 Db 50 TACTGTGTTTAAAGAGATGTTGTACCAAGTAACTAATCTATCGCTGT 1

RESULT 10  
 BE186786/c 467 bp mRNA linear EST 22-JUN-2000  
 LOCUS 946012C08.XI.946 - tassal primordialium prepared by schmidt lab Zea  
 DEFINITION mays cDNA, mRNA sequence.  
 ACCESSION BE186786  
 VERSION BE186786.1 GI:8665970  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 467)

REFERENCE 1 Walbot, V.  
 Maize ESTs from various cDNA libraries sequenced at Stanford  
 UNIVERSITY  
 Unpublished (1999)  
 CONTACT: Walbot V  
 JOURNAL Department of Biological Sciences  
 COMMENT Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 723 8221  
 Email: walbot@stanford.edu  
 Plate: 946012 row: C column: 08.  
 Location/Qualifiers

FEATURES  
 source  
 1..467  
 /organism="Zea mays"  
 /cultivar="OH43"  
 /db\_xref="taxon:4577"  
 /clone\_lib="946 - tassal primordialium prepared by Schmidt  
 lab"  
 /issue\_type="tassels"  
 /dev\_stage="just after the transition from vegetative to  
 inflorescence development"  
 /lab\_host="XL0LR"  
 /note="Organ: tassels; Vector: HybridZAP; Site\_1: EcoRI;  
 Site\_2: XhoI; George Chuck dissected immature tassels  
 between 1mm and 3mm. Sharon Stanfield prepared the cDNA  
 library in HybridZAP. Sample insert size range was 350 bp  
 to 3 Kb with a 1 Kb average."  
 110 c 96 g 153 t

BASE COUNT 108 a 110 c 96 g 153 t  
 ORIGIN  
 Query Match 30.6%; Score 448; DB 9; Length 467;  
 Best Local Similarity 98.7%; Pred. No. 1e-55; Mismatches 5; Indels 1; Gaps 1;  
 Matches 462; Conservative 0;

QY 886 AGATATCAATTCCTGAGAGCTGGCTTACCAAGAGCTCGAGCTTGTTCAGAGAGCCT 945  
 |||||||  
 Db 467 AGATATCAATTCCTGAGAGCTGGCTTACCAAGAGCTCGAGCTTGTTCAGAGAGCCT 408  
 |||||||  
 QY 946 AATGTGCATTTGATATTCCTGAGTAATAGACTGCACCTGATGAGAGAGCTCTATA 1005  
 |||||||  
 Db 407 AATGTGCATTTGATATTCCTGAGTAATAGACTGCACCTGATGAGAGAGCTCTATA 348  
 |||||||  
 QY 1006 AGTTTCTGTAAAGATAATGTTTCAACGAGATCGGCTGACAAAGCCATAGAGAG 1065  
 |||||||  
 Db 347 AGTTTCTGTAAAGATAATGTTTCAACGAGATCGGCTGACAAAGCCATAGAGAG 288  
 |||||||

QY 1066 ATCAATTCGCAAGATTAATGCTGCGAAGAGACTGAGTCTCTTTTCAAGCAACT 1125  
 |||||||  
 Db 287 ATCAATTCGCAAGATTAATGCTGCGAAGAGACTGAGTCTCTTTTCAAGCAACT 228  
 |||||||  
 QY 1126 GCCACACATGACGACCGCTTAAAGAGGAGACCTTCGATTAACAGCAAGCAGCT 1185  
 |||||||  
 Db 227 GCCACACATGACGACCGCTTAAAGAGGAGACCTTCGATTAACAGCAAGCAGCT 168  
 |||||||  
 QY 1186 GCGAACAAGAAACAAAGAGCTGTGGGAAGACAAATATCTTGGATGCTTATGTACAA 1245  
 |||||||  
 Db 167 GCGAACAAGAAACAAAGAGCTGTGGGAAGACAAATATCTTGGATGCTTATGTACAA 108  
 |||||||  
 QY 1246 CTAGACTACGAAAGCAGCGGTGGCTGATCATCTGCTTAGATTATTAATCCTGTT 1305  
 |||||||  
 Db 107 CTAGACTACGAAAGCAGCGGTGGCTGATCATCTGCTTAGATTATTAATCCTGTT 49  
 |||||||  
 QY 1306 TTAACTCAGAGCTTTGGTAAAGTTTCTCATGTTTCAAGCTGGGCTA 1353  
 |||||||  
 Db 48 TTAACTCAGAGCTTTGGTAAAGTTTCTCATGTTTCAAGCTGGGCTA 1

RESULT 11  
 AWS62517/c 475 bp mRNA linear EST 10-MAR-2000  
 LOCUS 660065H06.X2.660 - Mixed stages of anther and pollen Zea mays cDNA,  
 DEFINITION mRNA sequence.  
 ACCESSION AWS62517  
 VERSION AWS62517.1 GI:7216395  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 475)

REFERENCE 1 Walbot, V.  
 Maize ESTs from various cDNA libraries sequenced at Stanford  
 UNIVERSITY  
 Unpublished (1999)  
 CONTACT: Walbot V  
 JOURNAL Department of Biological Sciences  
 COMMENT Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 723 8221  
 Email: walbot@stanford.edu  
 Plate: 660065 row: H column: 06.  
 Location/Qualifiers

FEATURES  
 source  
 1..475  
 /organism="Zea mays"  
 /cultivar="OH43"  
 /db\_xref="taxon:4577"  
 /clone\_lib="660 - Mixed stages of anther and pollen"  
 /issue\_type="whole premeiotic anthers to pollen shed"  
 /dev\_stage="premeiotic anthers to pollen shed"  
 /lab\_host="XL0LR"  
 /note="Organ: anthers; Vector: Lambda Zap; Site\_1: EcoRI;  
 Site\_2: XhoI; Anther and pollen cDNA library.  
 Directionally sequenced with 5' end at the EcoRI site.  
 Created by Anne Franklin."

BASE COUNT 124 a 110 c 94 g 147 t  
 ORIGIN  
 Query Match 29.4%; Score 429.8; DB 9; Length 475;  
 Best Local Similarity 98.2%; Pred. No. 4.3e-53; Mismatches 7; Indels 1; Gaps 1;  
 Matches 445; Conservative 0;

QY 979 ACTGACCTGATGAGAGAGGCTCATAGATTCTCTGTAAAGATATGTTTCAACGAA 1038  
 |||||||  
 Db 475 ACTGACCTGATGAGAGAGGCTCATAGATTCTCTGTAAAGATATGTTTCAACGAA 416  
 |||||||  
 QY 1039 GATCGGTGACCAAGGCCATAGAGATCAATCTGCCAAGATAATGTCGCAAGGA 1098  
 |||||||

```

|||||
Db 415 GATCGGGGTACAAAGCCATAGAGATCAAAATCGCAAGAAATTAATGCTCCCAAGGA 356
QY 1099 AGACTGAGTCTCTTTTCAAGCCCACTCCACACATCAGCACCCTAAACGGAAG 1158
Db 355 AGACTGAGTCTCTTTTCAAGCCCACTCCACACATCAGCACCCTAAACGGAAG 296
QY 1159 ACTCGGATAAAAACAAGCAGCAGCTCGCAACAAGAAACAAAGGCTGGTGAAG 1218
Db 295 ACTTCGATAAAAACAAGCAGCAGCTCGCAACAAGAAACAAAGGCTGGTGAAG 236
QY 1219 AATATATCTTGATGCTTGATGTACACTACAGCTACGAAGCAGCGGTGCGATCAC 1278
Db 235 AATATATCTTGATGCTTGATGTACACTACAGCTACGAAGCAGCGGTGCGATCAC 176
QY 1279 TTGCTTAATATTTAACTCCCTGTTTAACTCAGACTTGGTAAAGCTTGCATG 1338
Db 175 TTGCTTAATATTTAACTCCCTGTTTAACTCAGACTTGGTAAAGCTTGCATG 116
QY 1339 TTTCAGCTGGGTAAGTGTGTGTGAAGAGATTGTACCAAGTAAACAACTT 1398
Db 115 TTTCAGCTGGGTAAGTGTGTGTGAAGAGATTGTGTACCAAGTAAACAACTT 57
QY 1399 ATGCTGTTTTTACTCTCTCTCTTTGAAGTA 1431
Db 56 ATGCTGTTTTTACTCTCTCTCTTTGAAGTA 24

```

```

RESULT 12
AM774700 733 bp mRNA linear EST 07-SEP-2000
LOCUS EST333851 KV3 Medicago truncatula cDNA clone PKV3-23L21. mRNA
DEFINITION sequence.
ACCESSION AM774700
VERSION AM774700.1 GI:7718617
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
REFERENCE 1 (bases 1 to 733)
AUTHORS Vandenbosch,K., Hurt,J., Moore,J., Beremand,P., Peng,H., Ellis,L.,
Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and
Fraser,C.M.
ESTs from roots of Medicago truncatula after Rhizobium inoculation
Unpublished (1999)
CONTACT: Vandenbosch K
DEPARTMENT: Department of Biology
INSTITUTION: Texas A&M University
LOCATION: College Station, TX 77843-3258, USA
TELEPHONE: Tel: 409 845 7707
FAX: Fax: 409 845 2891
EMAIL: Email: kate@tam.bio.tamu.edu
TEXT: Texas A&M EST name: T258119e
TIGR sequence name: MTEBE717K
More information is available at:
http://chryslie.tamu.edu/medicago
Seq primer: Skmod (CTA GAA CTA gtc gat CC).
Location/Qualifiers
1..733
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone_id="PKV3-23L21"
/clone_lib="KV3"
/tissue_type="Seedling roots"
/dev_stage="3 days post-inoculation with sinorhizobium
mellioli"
/lab_host="E. coli strain XL0LR"
/note="Vector: Bluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA

```

FEATURES

source

was directionally ligated into the Unizap XR vector from  
Stratagene and packaged using GigaPack IIT Gold packaging  
extracts. Plasmids containing cDNA inserts were excised  
from the recombinant lambda-Zap phage using Ex-assist  
helper phage and propagated in XL0LR cells."

```

BASE COUNT      222 a      125 c      183 g      203 t
ORIGIN
Query Match      28.7%; Score 420.4; DB 9; Length 733;
Best Local Similarity 75.4%; Pred. No. 8.3e-52;
Matches 523; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
QY 84 GATGGGATCAAGGGTTGAGCAAACTGTCGGCGAACAATGCCGCCAAGCGATGAAGA 143
Db 40 GATGGGATTAAGGGTTTAAGCAAGCTTTTACGTGATATGCTCCAAATGAATGAAGA 99
QY 144 GCAGAAATTCGAGAGCTCTTCGGCCGCAAAATGCCGCTGAGAGCCAGCATATATA 203
Db 100 GACAAATTCGAATCTTACCTTGGGCGTAAGATTGCTGTGATGCTGTATGAGCATTTA 159
QY 204 CCAGTCTCGATTGTAGTTCGACACAGCAGCATGGAACCTCTCAAAATGAAGCTGTGA 263
Db 160 CCAGTCTCTATTGTGTGGGAGAGAGTGAATGTGACTATATGATGAAGCTGTGA 219
QY 264 AATCACTAGTCAATTTGCAAGATGTTCAACCGACATTAAGATTACTGAGACGGGAAT 323
Db 220 ACTAATAGTCAATTTGCAAGATGTTGCGGACATCAAGCTCTTCAAGACCGGGAT 279
QY 324 CAAGCAGTTATGTTTTGATGCGACCTCTCGATATGAAGAAACAGAGCTCTCTAAT 383
Db 280 GAAGCAGATATATGTTTTGATGAGGAAAGCCACCGGATGAAGATCAAGACCTGAATA 339
QY 384 AAGATATCAAAAAGATGATGACACCAAGATCTGACTGAGGCACTTAAGGTAGAGA 443
Db 340 AGCTCTCTCAAGAGAGCTGAGGCTCGCGGTTTACAGACCTCTGATGAGGCTGACA 399
QY 444 TAAAGATGATGTAATAATTTAGCAGAGAGACTGTAAAGTCAACAAGCAACAAGCA 503
Db 400 TAAGAGATGATTAATAATTTAGCAGAGAGACTGTAAAGTCAACAAGCAACAAGCA 459
QY 504 AATTTGTAAGCGGATTAAGGCTTAAAGGCTTCTGTTGATGAGGACCTCTCTGAGC 563
Db 460 CCAGTCTCAAAAAGACTTTTGACAGCTCATGAGAGCTGCTGTTGAGGACCTCTGAGC 519
QY 564 AAGACAGAGATGAGAGCTTTTGACATTAAGATAGGCTGCTGCTTCTTCAAGAGA 623
Db 520 AAGAGCTGAGTGTCTGCTGACATTTGCAAGAGCTGAAAGGTGATGCTGCTCTCAAGA 579
QY 624 TATGACATCCCTTACTTTTGGGGCTCCAGGCTTCTGATTTAATGATGATGAGG 683
Db 580 CATGATTCCTTAACGCTTTGAGAGCTCTAAGTTCTTCCGCACTTAATGATGATGAG 639
QY 684 CAGAAATATCTGTGATGATGATTTGATGATGATGATGATGATGATGATGATGATG 743
Db 640 AAGAAAGATTCAGATTAATGATGATTTGGGCTTGAACAATTTGAGAGGCTAGATTGAC 699
QY 744 CAGGACAGATTCATGATTTGATGATGATGATGATGATGATGATGATGATGATG 777
Db 700 CCGGACCAATTTATGACTTATGATGATGATGATGATGATGATGATGATGATGATG 733

```

```

RESULT 13
AI834484/C 553 bp mRNA linear EST 02-FEB-2000
LOCUS 606068609.x1 606 - Ear tissue cDNA library from Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION AI834484
VERSION AI834484.1 GI:5468693
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

```

clade; Panicoidae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 553)

AUTHORS Malbot, V.

JOURNAL Maize ESTs from various cDNA libraries sequenced at Stanford

COMMENT University of California, Palo Alto, CA 94304, USA

unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 8221

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 606068 row: G column: 09.

Location/Qualifiers

1..553

/organism="Zea mays"

/cultivar="Oh1043"

/db\_xref="taxon:4577"

/clone\_lib="606 - Ear tissue cDNA library from Schmidt

lab"

/tissue\_type="mixed"

/dev\_stage="ear length from 0.5 cm - 2.0 cm"

/lab\_host="X10LR (Stratagene)"

/note="Organ: immature ear; Vector: PBK-CMV; Site\_1: EcoRI

; Site\_2: XhoI; Mixed ear tissue cDNA library from Schmidt

lab"

BASE COUNT 140 a 112 c 106 g 195 t

ORIGIN

Query Match 28.2%; Score 413; DB 9; Length 553;

Best Local Similarity 96.6%; Pred. No. 1.1e-50;

Matches 422; Conservative 0; Mismatches 15; Indels 0; Caps 0;

QY 995 AAGGCTCATAGTTCCTGCTAAAGATATGCTCAAGAGATGCGGAGCAAGG 1054

Db 553 AAGGCTCATAGTTCCTGCTAAAGATATGCTCAAGAGATGCGGAGCAAGG 494

QY 1055 CCATGAGAGAGATCAATCTGCCAAGATTAATCGTGCAGAGAGACTCGATCTTTT 1114

Db 493 CCATGAGAGAGATCAATCTGCCAAGATTAATCGTGCAGAGAGACTCGATCTTTT 434

QY 1115 TCAGAGCAAGTGCAGACATCAGACCGCTTAAGAGAGAGAGAGAGAGAGAGAG 1174

Db 433 TCAGAGCAAGTGCAGACATCAGACCGCTTAAGAGAGAGAGAGAGAGAGAGAG 374

QY 1175 GCAAGGCAAGTGCAGACATCAGACCGCTTAAGAGAGAGAGAGAGAGAGAGAG 1234

Db 373 GCAAGGCAAGTGCAGACATCAGACCGCTTAAGAGAGAGAGAGAGAGAGAGAG 314

QY 1235 TTGATGTACAACTACGACTACGAAAGCAGCGGTGATGATCACTTCGCTTATTT 1294

Db 313 TTGATGTACAACTACGACTACGAAAGCAGCGGTGATGATCACTTCGCTTATTT 254

QY 1295 AACCTCCGTTTAACTACGAGCTTTGTTAAAGTTCGCTTAAAGTTCGCTTAAAG 1354

Db 253 AACCTCCGTTTAACTACGAGCTTTGTTAAAGTTCGCTTAAAGTTCGCTTAAAG 194

QY 1355 GTTACTGTGTTTGAAGAGATGCTGACCAAGTAAACAAACTTATCGCTTTTACT 1414

Db 193 GTTACTGTGTTTGAAGAGATGCTGACCAAGTAAACAAACTTATCGCTTTTACT 134

QY 1415 TCTTGCTCTTGAAGTA 1431

Db 133 TCTTGCTCTTGAAGTA 117

RESULT 14

LOCUS BG263233 539 bp mRNA linear EST 16-FEB-2001

DEFINITION WHE2339\_F04.L0725 Wheat pre-anthesis spike cDNA library Triticum

aestivum cDNA clone WHE2339\_F04.L07, mRNA sequence.

ACCESSION BG263233

VERSION BG263233.1 GI:12865185

KEYWORDS EST.

SOURCE bread wheat.

ORGANISM Triticum aestivum

REFERENCE 1 (bases 1 to 539)

AUTHORS Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han

, P.S., Hsiao, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,

Seaton, C.L., and Tong, J.C.

THE structure and function of the expressed portion of the wheat

genomes - Pre-anthesis spike cDNA library

unpublished (2000)

CONTACT: Olin Anderson

US Department of Agriculture, Agricultural Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@w.usda.gov

Sequence have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20

Seq primer: Stratagene, SK primer.

Location/Qualifiers

1..539

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone\_lib="WHE2339\_F04.L07"

/tissue\_type="Spike before anthesis"

/dev\_stage="Adult plant"

/lab\_host="E. coli SOLR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;

Site\_1: EcoRI; Site\_2: XhoI; Plants were grown in the

greenhouse. Whole spike with awns trimmed, while, green

and yellow anther were collected and total RNA, and

poly(A) RNA were prepared, a cDNA library was made, and

the cDNA clones were in vivo excised to give phuscript

phagmids in the T3 Close lab (Choi, Close, Fenton) at

the University of California, Riverside. Plasmid DNA

preparations and DNA sequencing were performed in the OD

Anderson lab (all other authors)."

BASE COUNT 163 a 108 c 141 g 126 t 1 others

ORIGIN

Query Match 28.1%; Score 411.6; DB 10; Length 539.

Best Local Similarity 85.2%; Pred. No. 1.8e-50;

Matches 459; Conservative 0; Mismatches 80; Indels 0; Caps 0;

QY 160 TACTTCGCGCGCAAAATCGCGCTGCAGCGACATGATATACGAGTTCGATTGTA 219

Db 1 TACTTCGCGCGCGCAAAATCGCGCTGCAGCGACATGATATACGAGTTCGATTGTA 60

QY 220 GTTGAAGGAGCAGGCGATGGAACCTGCACAAATGAAGTGTGAAGTCACTAGTCAATTG 279

Db 61 GTTGAAGGAGCAGGCGATGGAACCTGCACAAATGAAGTGTGAAGTCACTAGTCAATTG 120

QY 280 CAAAGGATGTTCAACCGGACATTAACATTACTGGAAGCGGAGATCAAGCAATTATGTT 339

Db 121 CAAAGGATGTTCAACCGGACATTAACATTACTGGAAGCGGAGATCAAGCAATTATGTT 180

QY 340 TTGATGCGCAAGCGCTCTGATATGAAGAACAGAGCTTGTAAAGATCTCAAAAAGA 399

Db 181 TTGATGCGCAAGCGCTCTGATATGAAGAACAGAGCTTGTAAAGATCTCAAAAAGA 240

QY 400 GATGATCAACCAAGATCTGACTGAGCGAGTACGAGTACGATTAAGTTCGATTGAA 459

Db 241 AATTAACCAACAGAGCTGAGGAGAGAGTACGAGTACGAGTACGATTAAGTTCGATTGAA 300

QY 460 AAATTTGAGCAGAGGAGCTGTAAGGTCAACAAGCAACGAGAGATTTTAACCGGCTA 519

```

Db      301 AATTCAGCAGAGAACTTAAGTCAAGACAGACACATGATGTTTGAAGCTCTA 360
QY      520 TTAAAGCTATGGGGTTCCTTGTAGAGCACCCTTGTAGACAGACAGATGTGCA 579
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      361 CTAAAGCTGTGGTCTCTCTGTGAGCTCCCTGTGAGCAATATCATATGTGCT 420
QY      580 GCCCTTTCATAAAGCATTAAGGTTCCTGTGCTTCAGAAAGATATGAGCTCTTACT 639
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      421 GCCCTTTCAGAAATGACAAAGGTATGCTGTGATGCCAAGATATGAGCTCTTACT 480
QY      640 TTTGGGGCTCCACGCTTCCTCTGCTATTTAATGATCCCAAGTTCAGAAATACCTGT 698
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      481 TTTGAGCTACACGCTTGTCTGATTTGATGATTCAGATTCAGAAATACCTGT 539

RESULT 15
AM288784 414 bp mRNA linear EST 16-JAN-2000
LOCUS    707010C02.x5.707 - Mixed adult tissues from Walbot lab (SK) Zea
DEFINITION
mays cDNA, mRNA sequence.
ACCESSION
AM288784.1 GI:6695706
VERSION  AM288784.1
KEYWORDS
EST.
SOURCE   Zea mays.
ORGANISM Zea mays.
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
          clade; Panicoideae; Andropogoneae; Zea.
          1 (bases 1 to 414)
REFERENCE
  Walbot V.
  Maize ESTs from various cDNA libraries sequenced at Stanford
  University
  Unpublished (1999)
  Contact: Walbot V
  Department of Biological Sciences
  Stanford University
  855 California Ave, Palo Alto, CA 94304, USA
  Tel: 650 723 2227
  Fax: 650 725 8221
  Email: walbot@stanford.edu
  Plate: 707010 row: C column: 02.
  Location/Qualifiers
    source          1..414
                   /organism="Zea mays"
                   /cultivar="W23"
                   /db_xref="taxon:4577"
                   /clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
FEATURES
  source
    1..414
    /organism="Zea mays"
    /cultivar="W23"
    /db_xref="taxon:4577"
    /clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
    /tissue_type="tassel, kernel, silk, husk, root, leaf"
    /dev_stage="adult"
    /lab_host="PH10B"
    /note="Organ: tassel, kernel, silk, husk, root, leaf;
    vector: pGAD10; Site_1: EcoRI; cDNA library from fully
    differentiated maize tissues from an active Mutator
    plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
    husk, root, leaf). Unidirectionally cloned."
BASE COUNT      120 a      82 c      101 g      109 t      2 others
ORIGIN
Query Match      27.5%: Score 402.2; DB 9; Length 414;
Best Local Similarity 98.8%: Pred No. 4.4e-49;
Matches 404; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY      649 CCACGGTTCCTGCTCATTTAATGATCCAAAGTTCCTGATGAGTAATT 708
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      6  CCACGGTTCCTGCTCATTTAATGATCCAAAGTTCCTGATGAGTAATT 65
QY      709 GATGTGCCAAGGTTTGGAGGAGCTTGAAGTCAACATGAGACCATGATCATTTGTGC 768
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      66 GATGTGCCAAGGTTTGGAGGAGCTTGAAGTCAACATGAGACCATGATCATTTGTGC 125
QY      769 ATCTGTGTGATGTGACTATTGTGATACATCAAGATATCGGGGCAACAGCTCTG 828
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db      126 ATCTGTGTGATGTGACTATTGTATAGCATCAAGGATATCGGGGGCAACAGCTCTG 185
QY      829 AAATTAATTCGTCACATNGGTCATAGAAAGCATCTGAGAAATCTTAATAAGACAGA 888
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      186 AAATTAATTCGTCACATNGGTCATAGAAAGCATCTGAGAAATCTTAATAAGACAGA 245
QY      889 TATCAATTCCTGAGAGCTGCGCTTACCAGAAAGCTGAGCGTTGTTCAAGAGCTTAAT 948
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      246 TACCAATTCCTGAGAGCTGCGCTTACCAGAAAGCTGAGCGTTGTTCAAGAGCTTAAT 305
QY      949 GTACATTTGATATTCCTGAGCTTAAATGAGACCTGATGAGAGGCTCTCAATAAGT 1008
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      306 GTACATTTGATATTCCTGAGCTTAAATGAGACCTGATGAGAGGCTCTCAATAAGT 365
QY      1009 TTCTGTGTAAGATTAATGCTTTTCACGAGATCGGCTGACAAAGGCCA 1057
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      366 TTCTGTGTAAGATTAATGCTTTTCACGAGATCGGCTGACAAAGGCCA 414

```

Search completed: November 5, 2002, 13:36:44  
 Job time : 1271.78 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:23:34 : Search time 31 Seconds  
(without alignments)  
1357.967 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939  
Sequence: 1 MGIGLTKRLADNAPKAMKE.....SDTKRAAANKTKAGKKK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_032802.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1939	100.0	379	21	Maize Rad2/FEN-1 p
2	1939	100.0	379	21	Maize Rad2/FEN-1 p
3	1933	99.7	379	21	Maize Rad2/FEN-1 p
4	1933	99.7	379	21	Maize Rad2/FEN-1 p
5	1031	53.2	380	20	Maize Rad2/FEN-1 p
6	1007.5	52.0	377	20	Human FEN-1 protei
7	998.5	51.5	385	22	Mouse FEN-1 protei
8	996	51.4	378	20	Drosophila melanog
9	938	48.4	382	20	Human FEN-1 protei
10	699.5	36.1	373	22	Yeast FEN-1 protei
11	648.5	33.4	340	18	Novel human diagno
					Pyrococcus furiosu

12	648.5	33.4	340	19	AAW79970	Pyrococcus furiosu
13	648.5	33.4	340	19	AAW59940	Amino acid sequenc
14	647	33.4	326	19	AAW59953	Amino acid sequenc
15	644	33.2	332	19	AAW59951	Amino acid sequenc
16	636	32.8	343	22	AAW66644	Amino acid sequenc
17	632.5	32.6	343	22	AAW03778	Pyrococcus heat re
18	609	31.4	325	19	AAW59950	Amino acid sequenc
19	600	30.9	340	19	AAW59949	Amino acid sequenc
20	562	29.0	340	19	AAW59952	Amino acid sequenc
21	560.5	28.9	326	18	AAW24215	Methanococcus jann
22	560.5	28.9	326	18	AAW59939	Methanococcus jann
23	553.5	28.5	326	19	AAW79969	Amino acid sequenc
24	553.5	27.8	336	19	AAW79962	Methanococcus jann
25	538.5	27.8	336	19	AAW79962	Archaeoglobus fulg
26	525	27.1	328	19	AAW59946	Amino acid sequenc
27	411	21.2	328	19	AAW59948	Amino acid sequenc
28	340.5	17.6	258	19	AAW59947	Amino acid sequenc
29	336	17.3	1516	21	AAW2507	Yeast delta-RAD2 p
30	257	13.3	726	22	AAW18195	Plasmodium falcipla
31	243.5	12.6	1236	22	ABW65192	Drosophila melanog
32	243.5	12.6	1236	22	ABW65287	Drosophila melanog
33	210	10.8	732	22	ABW67273	Drosophila melanog
34	208	10.7	872	20	AAW31815	Thermococcus litor
35	205	10.6	872	18	AAW26605	Thermococcus litor
36	205	10.6	872	20	AAW31812	Thermococcus litor
37	205	10.6	872	20	AAW31816	Thermococcus litor
38	184.5	9.5	291	20	AAW78466	T. thermophilus DNA
39	181	9.3	832	13	AAW22602	Taq polymerase enc
40	174.5	9.0	832	13	AAW22603	Taq polymerase enc
41	174.5	9.0	832	13	AAW22604	Taq polymerase enc
42	173.5	8.9	834	17	AAW9542	Thermus flavus DNA
43	173.5	8.9	834	17	AAW9542	DNA polymerase I h
44	173	8.9	548	17	AAW6263	Mutant Thermus aqu
45	173	8.9	695	17	AAW6264	Mutant Thermus aqu

#### ALIGNMENTS

RESULT 1  
ID AAY95307 standard; protein: 379 AA.  
AC AAY95307;  
DT 12-SEP-2000 (first entry)  
XX  
DE Maize Rad2/FEN-1 protein.  
XX  
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;  
KW endonuclease; exonuclease; DNA repair; gene targeting.  
XX  
OS Zea mays.  
XX  
PN WO200036109-A1.  
XX  
PD 22-JUN-2000.  
XX  
PF 16-NOV-1999; 99WO-US27147.  
XX  
PR 15-DEC-1998; 98US-0112332.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Mahajan PB;  
XX  
DR MPI. 2000-452026/39.  
DR N-PSDB: AAA27923.  
XX  
PT Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA  
PT recombination and repair in transgenic plants, e.g. for gene targeting  
PT and the production of male sterile plants -

PS Claim 11: Page 71-72: 85bp: English.

CC The present sequence is that of maize Rad2/FEN-1, as deduced from  
CC a cDNA clone (see AAA27923) isolated from maize line B73 immature  
CC ear tissue. Rad2/FEN-1 is a structure specific endonuclease which  
CC under certain conditions also acts as an exonuclease. Rad2/FEN-1  
CC can be expressed in transgenic plant cells using conventional  
CC methods. The protein is involved in the regulation of DNA repair  
CC and recombination in plant systems and therefore may be used for  
CC improving gene targeting during further recombinant DNA protocols  
CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential  
CC in DNA replication and nucleotide excision and repair reactions.  
CC The exolytic activity is involved in double strand break repair and  
CC end joining. The protein is also useful in strand exchange  
CC reactions during homologous recombination. These functions may be  
CC useful in gene targeting and in the production of male sterile  
CC plants. The efficacy of gene targeting can be improved by the  
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
CC be produced by the down regulation of Rad2/FEN-1 expression.

SO Sequence 379 AA:

Query Match 100.0%; Score 1939; DB 21; Length 379;  
Best Local Similarity 100.0%; Pred. No. 1.2e-170;  
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGTKGLTKLLADNAPKAMKQKFESEFGKRIAVDASMSIYQFLIVGRTGMETLTNAGE 60  
DB 1 MGTKGLTKLLADNAPKAMKQKFESEFGKRIAVDASMSIYQFLIVGRTGMETLTNAGE 60  
QY 61 VTSHLQGMFNRTIRLLEAGIKPYVVDGKPPDMKKQELAKRYSKRDATKDLTEAVEYGD 120  
DB 61 VTSHLQGMFNRTIRLLEAGIKPYVVDGKPPDMKKQELAKRYSKRDATKDLTEAVEYGD 120  
QY 121 KDAIEKLSKRTYKVTROHNECDKRLRLMGVVPVPAFSAEACALCINIKVAVASSED 180  
DB 121 KDAIEKLSKRTYKVTROHNECDKRLRLMGVVPVPAFSAEACALCINIKVAVASSED 180  
QY 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAVLELELTMDQFLDCLTCGCCDYCSIK 240  
DB 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAVLELELTMDQFLDCLTCGCCDYCSIK 240  
QY 241 GIGGQIALKLIROHGSIESILENLNDRQIPEDMPYQEARLFEKPNVTLDIPELKWTA 300  
DB 241 GIGGQIALKLIROHGSIESILENLNDRQIPEDMPYQEARLFEKPNVTLDIPELKWTA 300  
QY 301 PDEEGLISFLVKNNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
DB 301 PDEEGLISFLVKNNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
QY 361 DKTSKAANKTKYAGGKK 379  
DB 361 DKTSKAANKTKYAGGKK 379

RESULT 2

AA95309 ID AAY95309 standard; Protein: 379 AA.

XX AAY95309;

DT 12-SEP-2000 (first entry)

DE Maize Rad2/FEN-1 protein.

KM Maize; Rad2/FEN-1; transgenic plant; male sterile plant;  
KW endonuclease; exonuclease; DNA repair; gene targeting.

XX Zea mays.

PN W0200036109-A1.

XX 22-JUN-2000.

XX 16-NOV-1999; 99WO-US27147.

XX 15-DEC-1998; 98US-0112332.

XX (PION-) PIONEER HI-BRED INT INC.

PI Mahajan PB;

DR WPI: 2000-452026/39.

XX N-PSDB: AAA27925.

PT Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA  
PT recombination and repair in transgenic plants, e.g. for gene targeting  
PT and the production of male sterile plants.

PS Example 1: Page 76-77: 85bp: English.

CC The present sequence is that of maize Rad2/FEN-1, as deduced from  
CC a cDNA clone (see AAA27925) derived from maize line W23 tassel  
CC polyA RNA. Rad2/FEN-1 is a structure specific endonuclease which  
CC under certain conditions also acts as an exonuclease. Rad2/FEN-1  
CC can be expressed in transgenic plant cells using conventional  
CC methods. The protein is involved in the regulation of DNA repair  
CC and recombination in plant systems and therefore may be used for  
CC improving gene targeting during further recombinant DNA protocols  
CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential  
CC in DNA replication and nucleotide excision and repair reactions.  
CC The exolytic activity is involved in double strand break repair and  
CC end joining. The protein is also useful in strand exchange  
CC reactions during homologous recombination. These functions may be  
CC useful in gene targeting and in the production of male sterile  
CC plants. The efficacy of gene targeting can be improved by the  
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
CC be produced by the down regulation of Rad2/FEN-1 expression.

SO Sequence 379 AA:

Query Match 100.0%; Score 1939; DB 21; Length 379;  
Best Local Similarity 100.0%; Pred. No. 1.2e-170;  
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGTKGLTKLLADNAPKAMKQKFESEFGKRIAVDASMSIYQFLIVGRTGMETLTNAGE 60  
DB 1 MGTKGLTKLLADNAPKAMKQKFESEFGKRIAVDASMSIYQFLIVGRTGMETLTNAGE 60  
QY 61 VTSHLQGMFNRTIRLLEAGIKPYVVDGKPPDMKKQELAKRYSKRDATKDLTEAVEYGD 120  
DB 61 VTSHLQGMFNRTIRLLEAGIKPYVVDGKPPDMKKQELAKRYSKRDATKDLTEAVEYGD 120  
QY 121 KDAIEKLSKRTYKVTROHNECDKRLRLMGVVPVPAFSAEACALCINIKVAVASSED 180  
DB 121 KDAIEKLSKRTYKVTROHNECDKRLRLMGVVPVPAFSAEACALCINIKVAVASSED 180  
QY 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAVLELELTMDQFLDCLTCGCCDYCSIK 240  
DB 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAVLELELTMDQFLDCLTCGCCDYCSIK 240  
QY 241 GIGGQIALKLIROHGSIESILENLNDRQIPEDMPYQEARLFEKPNVTLDIPELKWTA 300  
DB 241 GIGGQIALKLIROHGSIESILENLNDRQIPEDMPYQEARLFEKPNVTLDIPELKWTA 300  
QY 301 PDEEGLISFLVKNNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
DB 301 PDEEGLISFLVKNNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
QY 361 DKTSKAANKTKYAGGKK 379  
DB 361 DKTSKAANKTKYAGGKK 379

RESULT 3

AA95308

```

ID  AAY95308 standard; Protein; 379 AA.
XX
AC  AAY95308;
XX
DT  12-SEP-2000 (first entry)
XX
DE  Maize Rad2/FEN-1 protein.
XX
KM  Maize: Rad2/FEN-1; transgenic plant; male sterile plant;
XX  endonuclease; exonuclease; DNA repair; gene targeting.
XX
OS  Zea mays.
XX  WO200036109-A1.
XX  PD  22-JUN-2000.
XX
PF  16-NOV-1999; 99WO-US27147.
XX
PR  15-DEC-1998; 98US-0112332.
XX
PA  (PION-1) PIONEER HI-BRED INT INC.
XX
PI  Mahajan PB;
XX
DR  WPI: 2000-452026/39.
XX  N-PSDB; AAA27924.
XX
PS  Example 1; Page 74; 85pp; English.
XX
CC  The present sequence is that of maize Rad2/FEN-1, as deduced from
CC  a cDNA clone (see AAA27924) derived from maize line B73 seedling
CC  tissue RNA. Rad2/FEN-1 is a structure specific endonuclease which
CC  can be expressed in transgenic plant cells using conventional
CC  methods. The protein is involved in the regulation of DNA repair
CC  and recombination in plant systems and therefore may be used for
CC  improving gene targeting during further recombinant DNA protocols
CC  involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC  in DNA replication and nucleotide excision and repair reactions.
CC  The exolytic activity is involved in double strand break repair and
CC  end joining. The protein is also useful in strand exchange
CC  reactions during homologous recombination. These functions may be
CC  useful in gene targeting and in the production of male sterile
CC  plants. The efficacy of gene targeting can be improved by the
CC  overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC  be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ  Sequence 379 AA;

Query Match 99.7%; Score 1933; DB 21; Length 379;
Best Local Similarity 99.7%; Pred. No. 4.2e-170;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY  1 MGIKLTLLADNAPKAMKEGFESYFGRIADVASKMSYQVFIIVGRGMEITNEAGE 60
DB  1 MGIKLTLLADNAPKAMKEGFESYFGRIADVASKMSYQVFIIVGRGMEITNEAGE 60
OY  61 VTSHQGMFNRTIRLENGITPVYVFDGKPPDMKKOFLARXSKRDATKDLTEAVEVD 120
DB  61 VTSHQGMFNRTIRLENGITPVYVFDGKPPDMKKOFLARXSKRDATKDLTEAVEVD 120
OY  121 KDAIEKLSKRTVKYTRQHNEDEKRLRLMGVYVVEAPSEAEACALCINDKFAVASD 180
DB  121 KDAIEKLSKRTVKYTRQHNEDEKRLRLMGVYVVEAPSEAEACALCINDKFAVASD 180
OY  181 MDSLTFEGAPRFLRLHMDSSKKIPVMEFDVAKVLEELTMDQFIDICTICGDCYDSIK 240
DB  181 MDSLTFEGAPRFLRLHMDSSKKIPVMEFDVAKVLEELTMDQFIDICTICGDCYDSIK 240

```

```

OY  241 GIGGOTALKLIRQHSIESILENLKDRYQIPEDWPYQOEARLREKPNVTLDIPELKMTA 300
DB  241 GIGGOTALKLIRQHSIESILENLKDRYQIPEDWPYQOEARLREKPNVTLDIPELKMTA 300
OY  301 PDEEGLISFLVNDNGFNEDRVRKATEKIKSAKNKSSQGRLESPFKPATYSAPLKRRETS 360
DB  301 PDEEGLISFLVNDNGFNEDRVRKATEKIKSAKNKSSQGRLESPFKPATYSAPLKRRETS 360
OY  361 DKTSKAANKKTRKAGGKK 379
DB  361 DKTSKAANKKTRKAGGKK 379

RESULT 4
AAY95310
ID  AAY95310 standard; Protein; 379 AA.
XX
AC  AAY95310;
XX
DT  12-SEP-2000 (first entry)
XX
DE  Maize Rad2/FEN-1 protein.
XX
KM  Maize: Rad2/FEN-1; transgenic plant; male sterile plant;
XX  endonuclease; exonuclease; DNA repair; gene targeting.
XX
OS  Zea mays.
XX  WO200036109-A1.
XX  PD  22-JUN-2000.
XX
PF  16-NOV-1999; 99WO-US27147.
XX
PR  15-DEC-1998; 98US-0112332.
XX
PA  (PION-1) PIONEER HI-BRED INT INC.
XX
PI  Mahajan PB;
XX
DR  WPI: 2000-452026/39.
XX  N-PSDB; AAA27926.
XX
PS  Example 1; Page 79-80; 85pp; English.
XX
CC  The present sequence is that of maize Rad2/FEN-1, as deduced from
CC  a cDNA clone (see AAA27926) derived from maize line B73 endosperm
CC  RNA. Rad2/FEN-1 is a structure specific endonuclease which
CC  under certain conditions also acts as an exonuclease. Rad2/FEN-1
CC  can be expressed in transgenic plant cells using conventional
CC  methods. The protein is involved in the regulation of DNA repair
CC  and recombination in plant systems and therefore may be used for
CC  improving gene targeting during further recombinant DNA protocols
CC  involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC  in DNA replication and nucleotide excision and repair reactions.
CC  The exolytic activity is involved in double strand break repair and
CC  end joining. The protein is also useful in strand exchange
CC  reactions during homologous recombination. These functions may be
CC  useful in gene targeting and in the production of male sterile
CC  plants. The efficacy of gene targeting can be improved by the
CC  overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC  be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ  Sequence 379 AA;

Query Match 99.7%; Score 1933; DB 21; Length 379;
Best Local Similarity 99.7%; Pred. No. 4.2e-170;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 MGIGLTKLLADNAKRAKKEOKFESYFGKRIADVASMSTVOFLIVGRIGMETLTINEAGE 60
    |||||
DB 1 MGIGLTKLLADNAKRAKKEOKFESYFGKRIADVASMSTVOFLIVGRIGMETLTINEAGE 60
OY 61 VTSHLOGMNRITRLLENGIKPVYVFDGKPPDMKKOELARYSKRDPATKDLTEAVEVD 120
    |||||
DB 61 VTSHLOGMNRITRLLENGIKPVYVFDGKPPDMKKOELARYSKRDPATKDLTEAVEVD 120
OY 121 KDAIEKLSRTYKVTROHNEDCKRLRLMGVYVVEAPSEAEACALCINDKVFVASED 180
    |||||
DB 121 KDAIEKLSRTYKVTROHNEDCKRLRLMGVYVVEAPSEAEACALCINDKVFVASED 180
OY 181 MDSITFGARFLRLHMDPSKKIPYMEFDVAKVLEELTMDQFIDLCICGCDYCDISK 240
    |||||
DB 181 MDSITFGARFLRLHMDPSKKIPYMEFDVAKVLEELTMDQFIDLCICGCDYCDISK 240
OY 241 GIGGOTALKLIRHOGSISITENLNKORYQIPEDMPYOEARRLFEKEDNVTLDIPELKWT 300
    |||||
DB 241 GIGGOTALKLIRHOGSISITENLNKORYQIPEDMPYOEARRLFEKEDNVTLDIPELKWT 300
OY 301 PDEGLISFLVNDGPNEDRYTKAETIKSAKNSOGRLSEFPKPTATTSAPLRKETS 360
    |||||
DB 301 PDEGLISFLVNDGPNEDRYTKAETIKSAKNSOGRLSEFPKPTATTSAPLRKETS 360
OY 361 DKTSKAANKKTKAGKKK 379
    |||||
DB 361 DKTSKAANKKTKAGKKK 379

RESULT 5
AAM92504
ID AAM92504 standard; Protein: 380 AA.
AC AAM92504;
XX
DT 23-APR-1999 (first entry)
XX
DE Human FEN-1 protein.
XX
KW FEN-1; human; flap endonuclease; detection: diagnosis: carcinogen;
KW neoplasma; antineoplastic agent; cleavage.
XX
OS Homo sapiens.
XX
PN US5874283-A.
XX
PD 23-FEB-1999.
XX
PE 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
PI Harrington JJ, Hsieh C, Lieber MR;
XX
XX WPI; 1999-179985/15.
XX
XX N-PSDB; AAX02107.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing
XX e.g. recombinant polypeptides
XX
PS Claim 1; Fig 1A; 58pp: English.
XX
CC This sequence represents a human FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel

```

```

CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.
XX
SQ Sequence 380 AA;
Query Match 53.2%; Score 1031; DB 20; Length 380;
Best local Similarity 53.8%; Pred. No. 9.6e-87;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;
OY 1 MGIGLTKLLADNAKRAKKEOKFESYFGKRIADVASMSTVOFLIVGRIGMETLTINEAGE 60
    |||||
DB 1 MGIGLTKLLADNAKRAKKEOKFESYFGKRIADVASMSTVOFLIVGRIGMETLTINEAGE 60
OY 61 VTSHLOGMNRITRLLENGIKPVYVFDGKPPDMKKOELARYSKRDPATKDLTEAVEVD 120
    |||||
DB 61 VTSHLOGMNRITRLLENGIKPVYVFDGKPPDMKKOELARYSKRDPATKDLTEAVEVD 120
OY 121 KDAIEKLSRTYKVTROHNEDCKRLRLMGVYVVEAPSEAEACALCINDKVFVASED 180
    |||||
DB 121 KDAIEKLSRTYKVTROHNEDCKRLRLMGVYVVEAPSEAEACALCINDKVFVASED 180
OY 181 MDSITFGARFLRLHMDPSKKIPYMEFDVAKVLEELTMDQFIDLCICGCDYCDISK 240
    |||||
DB 181 MDSITFGARFLRLHMDPSKKIPYMEFDVAKVLEELTMDQFIDLCICGCDYCDISK 240
OY 241 GIGGOTALKLIRHOGSISITENLNKORYQIPEDMPYOEARRLFEKEDNVTLDIPELKWT 300
    |||||
DB 241 GIGGOTALKLIRHOGSISITENLNKORYQIPEDMPYOEARRLFEKEDNVTLDIPELKWT 300
OY 301 PDEGLISFLVNDGPNEDRYTKAETIKSAKNSOGRLSEFPKPTATTSAPLRKETS 360
    |||||
DB 301 PDEGLISFLVNDGPNEDRYTKAETIKSAKNSOGRLSEFPKPTATTSAPLRKETS 360
OY 359 TSDK--TSKAANKKTKAGKKK 379
    |||||
DB 358 PEPRGSTRKKKA--KTGAAGKFK 377

RESULT 6
AAM92505
ID AAM92505 standard; Protein: 377 AA.
AC AAM92505;
XX
DT 23-APR-1999 (first entry)
XX
DE Mouse FEN-1 protein.
XX
KW FEN-1; mouse; flap endonuclease; detection: diagnosis: carcinogen;
KW neoplasia; antineoplastic agent; cleavage.
XX
OS Mus sp.
XX
PN US5874283-A.
XX
PD 23-FEB-1999.
XX
PE 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
PI Harrington JJ, Hsieh C, Lieber MR;
XX
XX WPI; 1999-179985/15.
XX
XX N-PSDB; AAX02108.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing

```

PT e.g. recombinant polypeptides  
 XX  
 PS Claim 1; Fig 2A; 58pp; English.  
 XX  
 CC This sequence represents a mouse FEN-1 (flap endonuclease) protein. This  
 CC protein can be used in methods for detecting a pathological condition in  
 CC a patient, for diagnostic purposes, for screening for antineoplastic  
 CC agents and carcinogens, for diagnostic staging of neoplasms, for  
 CC producing recombinant flap endonuclease for use as research or  
 CC diagnostic reagents, for producing antibodies reactive with the novel  
 CC polypeptides, for producing transgenic nonhuman animals expressing the  
 CC novel polypeptides encoded by a transgene. The invention also provides  
 CC novel molecular cloning techniques and reagents involving cleavage of  
 CC a flap or nick with a flap endonuclease.  
 CC  
 XX  
 SO Sequence 377 AA;

Query Match 52.0%; Score 1007.5; DB 20; Length 377;  
 Best Local Similarity 53.4%; Pred. No. 1.4e-84;  
 Matches 203; Conservative 68; Mismatches 100; Indels 9; Gaps 7;  
 QY 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 60  
 DB 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 59  
 QY 61 VTSHLQGMENRTIRLEAGIKRPYVFDGKPPMKKQELAKRYSKRDADTKDLTEAVEVGD 120  
 DB 60 TTS-LMGFRTIR-MENGIKRPYVFDGKPPMKKQELAKRYSKRDADTKDLTEAVEVGD 116  
 QY 121 KDAIEKLSKRTVAVRQHNEDCKRLRLMGVYVEAPEAECAALCINDKFAVAASED 180  
 DB 117 MEEVEKTKRLVAVRQHNEDCKRLRLMGVYVEAPEAECAALCINDKFAVAASED 176  
 QY 181 MDSLTFGAPRLRHLMDPPSKKIPYMEPDVAKVLEBELTMDQFLDCLIGCDVCDYSIK 240  
 DB 177 MDCLTFSPVLMHLTASKEKKIPIDIEFHLSVYLOELIGNQEPYDCLILGSDYCESIR 236  
 QY 241 GIGGQATLKLIRQHSIESILENLKDRYQIPEDMYPQEARLFEKPNVT-LDIPELKWT 299  
 DB 237 GIGKRAVVDLIQHKSIIEELVRLDPSKYVPENMLHKEAQLLEPEVDPESVELKWS 296  
 QY 300 APDEGLISFLVNGFENDRYTKAIEKISAKNKSOGRLSEFPKPTATTSAPLKRKET 359  
 DB 297 EPNEELVTKMCEKQPSERIRISGVKRLSKSGSTQGRLDLDFKVGYSLS-AKREP 355  
 QY 360 SDRKTSKAANKKTKAGKK 379  
 DB 356 EPKGS--AKKAKGTGAGK 372

RESULT 7  
 ABB63960  
 ID ABB63960 standard; Protein; 385 AA.  
 XX  
 AC ABB63960;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 18672.  
 XX  
 KW Drosophila: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN W0200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001MO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.

XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Ventler JC, Adams M, Li FWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 XX  
 N-PSDB; ABL08063.  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 18672; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130311), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (AB57737-AB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
 XX  
 SO Sequence 385 AA;

Query Match 51.5%; Score 998.5; DB 22; Length 385;  
 Best Local Similarity 52.9%; Pred. No. 9.9e-84;  
 Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps 5;  
 QY 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 60  
 DB 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 59  
 QY 61 VTSHLQGMENRTIRLEAGIKRPYVFDGKPPMKKQELAKRYSKRDADTKDLTEAVEVGD 120  
 DB 60 PTHLMGMFRTIRLDNGIRKPYVFDGKPPDLKSGELAKRERREAEKALKAADAD 119  
 QY 121 KDAIEKLSKRTVAVRQHNEDCKRLRLMGVYVEAPEAECAALCINDKFAVAASED 180  
 DB 120 DAGIEFNNRLVAVRQHNEDCKRLRLMGVYVEAPEAECAALCINDKFAVAASED 179  
 QY 181 MDSLTFGAPRLRHLMDPPSKKIPYMEPDVAKVLEBELTMDQFLDCLIGCDVCDYSIK 240  
 DB 180 MDALTFGSTRILRYLTYSKARKMPYKESYDKLLEGLAINNREFIDICILGCDYCESIR 239  
 QY 241 GIGGQATLKLIRQHSIESILENLKDRYQIPEDMYPQEARLFEKPNVT-LDIPELKWT 299  
 DB 240 GIGKRAIETLINTYRDIELTLDNLDSSKYVPENMNVKVAARELJFPEVADADSIDLKW 299  
 QY 300 APDEGLISFLVNGFENDRYTKAIEKISAKNKSOGRLSEFPKPTATTSAPLKRKET 354  
 DB 300 EPDEGLVKEFLGDRQFNEERVRNGAKLMSKQAOVQRLDSEFKTLPSTPNATNA-- 357  
 QY 355 KRKETSDRKTSKAANKKTKAGG 376  
 DB 358 KKK--AEAKKSAANKKAKTSG 377

RESULT 8  
 AAW92508  
 ID AAW92508 standard; Protein; 378 AA.  
 XX  
 AC AAW92508;  
 XX  
 DT 23-APR-1999 (first entry)  
 XX  
 DE Human FEN-1 protein.  
 XX  
 KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;  
 KW neoplasma; antineoplastic agent; cleavage.

XX Homo sapiens.  
 OS  
 XX US5874283-A.  
 PN  
 XX 23-FEB-1999.  
 PD  
 XX 30-MAY-1995; 9505-0455968.  
 PF  
 XX 30-MAY-1995; 9505-0455968.  
 PR  
 XX 30-MAY-1995; 9505-0455968.  
 PA  
 XX (HARR/) HARRINGTON J J.  
 PA (HSTE/) HSIEH C.  
 PA (LIEB/) LIEBER M R.  
 XX Harrington JJ, Hsieh C, Lieber MR;  
 PI  
 XX WPI; 1999-179985/15.  
 DR  
 XX N-PSDB; AAX02111.  
 PT  
 XX DNA encoding flap endonuclease polypeptides - useful for producing  
 PS e.g. recombinant polypeptides  
 XX  
 PS Disclosure; Fig 5A-B; 58pp; English.  
 XX  
 CC This sequence represents a human FEN-1 (flap endonuclease) protein. This  
 CC protein can be used in methods for detecting a pathological condition in  
 CC a patient, for diagnostic purposes, for screening for antineoplastic  
 CC agents and carcinogens, for diagnostic staging of neoplasia, for  
 CC producing recombinant flap endonuclease for use as research or  
 CC diagnostic reagents, for producing transgenic nonhuman animals expressing the  
 CC novel polypeptides, for producing transgenic nonhuman animals expressing the  
 CC novel polypeptides encoded by a transgene. The invention also provides  
 CC novel molecular cloning techniques and reagents involving cleavage of  
 CC a flap or nick with a flap endonuclease.  
 CC  
 XX  
 XX Sequence 378 AA:

Query Match 51.4%; Score 996; DB 20; Length 378;  
 Best Local Similarity 52.4%; Pred. No. 1,6e-83;  
 Matches 199; Conservative 70; Mismatches 103; Indels 8; Gaps 6;

QY 1 MGKGLTKLLADNAPKAMEQKFESEYFGKRIAVDASMSIYOFLLVYGTGNETLTNEAGE 60  
 DB 1 MEHGLAKLLADVAPSAIRNDIKSYFGKRYAIDASMSIYQFLIAV-RGGCVILNNEGE 59  
 QY 61 VTSHLQGMFNRTIRLEAGIKPYVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEYGD 120  
 DB 60 TTS-LMGMEYRTMK-MENGIKPYVYFDGKPPQLKSGELAKRSRRRAEAKQLOQAOQAGM 117  
 QY 121 KDAIEKLSKRYKVTROHNEDECKRLRLMGVYVVEAPSEAEACALCINDKVFVASED 180  
 DB 118 EEEVEKTKRYKVTROHNEDECKRLRLMGVYVVEAPSEAEACALCINDKVFVASED 177  
 QY 181 MDSLTFGAPFLRLHMDPSSKKIPVMEFDVAKVLEELLETMDQFLDCTLCGCCDYSIR 240  
 DB 178 MDCLTFSPVYRLHMLTSEAKKLPIDQPHLSRYLQELGNOGFVDLCILSDVCESIR 237  
 QY 241 GIGGOTALKLIROHGSIESILENLNDROYIPEDWPYQEARLFEKPNVT-LDITELKWT 299  
 DB 238 GIGKRAVDLQKHSIEELIVRLDPSKYVPENMLAKRAQQLFEPEVADSEVELKWS 297  
 QY 300 APDEGLISFLVKNQNGFNEDRVTKATEIKSARKNSQGRLESEFFPTATTSAPLRKET 359  
 DB 298 EPNEELVKFKGCEKOFEEERIRISGVKRLSKNSQSGTQGRLDDEFFVIGTSSS-AKRREP 356  
 QY 360 SDKTSKAANKKTKAGGKK 379  
 DB 357 E---PKGPARKKAKTGAGK 373

RESULT 9  
 AAW92506

ID AAW92506 standard; Protein; 382 AA.  
 XX  
 AC AAW92506;  
 XX  
 DT 23-APR-1999 (first entry)  
 XX  
 DE Yeast FEN-1 protein.  
 XX  
 KW FEN-1; yeast; flap endonuclease; detection; diagnosis; carcinogen;  
 KW neoplasia; antineoplastic agent; cleavage.  
 OS Saccharomyces cerevisiae.  
 XX  
 PN US5874283-A.  
 XX  
 PD 23-FEB-1999.  
 XX  
 PF 30-MAY-1995; 9505-0455968.  
 XX  
 PR 30-MAY-1995; 9505-0455968.  
 XX  
 PA (HARR/) HARRINGTON J J.  
 PA (HSTE/) HSIEH C.  
 PA (LIEB/) LIEBER M R.  
 XX Harrington JJ, Hsieh C, Lieber MR;  
 PI  
 XX WPI; 1999-179985/15.  
 DR  
 XX N-PSDB; AAX02109.  
 PT  
 XX DNA encoding flap endonuclease polypeptides - useful for producing  
 PS e.g. recombinant polypeptides  
 XX  
 PS Disclosure; Fig 3A; 58pp; English.  
 XX  
 CC This sequence represents a yeast FEN-1 (flap endonuclease) protein. This  
 CC protein is used in a method to isolate novel human FEN-1 proteins for  
 CC detecting a pathological condition in a patient, for diagnostic purposes,  
 CC for screening for antineoplastic agents and carcinogens, for diagnostic  
 CC staging of neoplasia, for producing recombinant flap endonuclease for use  
 CC as research or diagnostic reagents, for producing transgenic nonhuman animals  
 CC with the novel polypeptides, for producing transgenic nonhuman animals  
 CC expressing the novel polypeptides encoded by a transgene. The invention  
 CC also provides novel molecular cloning techniques and reagents involving  
 CC cleavage of a flap or nick with a flap endonuclease.  
 CC  
 XX  
 XX Sequence 382 AA:

Query Match 48.4%; Score 938; DB 20; Length 382;  
 Best Local Similarity 49.5%; Pred. No. 3,8e-78;  
 Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

QY 1 MGKGLTKLLADNAPKAMEQKFESEYFGKRIAVDASMSIYOFLLVYGTGNETLTNEAGE 60  
 DB 1 MGKGINALISIEHPVAPKIRKSDIKSEFGKRYAIDASMSIYQFLIAVRQODGQULTNEAGE 60  
 QY 61 VTSHLQGMFNRTIRLEAGIKPYVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEYGD 120  
 DB 61 TTS-LMGMEYRTMK-MENGIKPYVYFDGKPPQLKSGELAKRSRRRAEAKQLOQAOQAGM 117  
 QY 121 KDAIEKLSKRYKVTROHNEDECKRLRLMGVYVVEAPSEAEACALCINDKVFVASED 180  
 DB 121 K--MQOERLVYVSKENHEEAQKLLGTMGIPYITAPTEAEQAQELAKKGVYAAASED 177  
 QY 181 MDSLTFGAPFLRLHMDPSSKKIPVMEFDVAKVLEELLETMDQFLDCTLCGCCDYSIR 240  
 DB 178 MDCLTFSPVYRLHMLTSEAKKLPIDQPHLSRYLQELGNOGFVDLCILSDVCESIR 237  
 QY 241 GIGGOTALKLIROHGSIESILENLNDROYIPEDWPYQEARLFEKPNVTLDIPE 295  
 DB 238 GIGVATLAKLIKTHGSIKIVFIEGSGESNNTKMKIPEDWPYQARMLFDPEV-IDGNE 296  
 QY 296 --LKWTPADEGLISFLVKNQNGFNEDRVTKATEIKSARKNSQGRLESEFFPTATT--- 350

```
DB 297 INKMSPPKREKLEIYLCDDKFESERYKSGISRLKGLKSGIGRLDGFQVVPKTRKEQ 356
QY 351 -SAPLKRKETSQKTSKAANKTKR 373
DB 357 LAAAKRAOENKRLNK-NKNKYTK 379

RESULT 10
ABG19545
ID ABG19545 standard; Protein: 373 AA.
AC ABG19545;
XX
XX
XX 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #19536.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-0508631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX N-PSDB: AAS83732.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity
XX
XX Claim 20: SEQ ID No 49904; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 373 AA;
XX
XX Query Match 36.1%; Score 699.5; DB 22; Length 373;
XX Best Local Similarity 41.1%; Pred. No. 4.1e-56;
XX Matches 155; Conservative 68; Mismatches 105; Indels 49; Gaps 7;
```

```
QY 1 MGKGTETLADNMPKAMEKQEFSEYGRKIAVDASMTIYQILIVGRIGMETLTNEAGE 60
DB 36 MGIGLAKLIADVAPSAIRENDIKSYFERKVAIDASMTIHFLLAV-RGGGDLVQNEEGE 94
QY 61 VTSHLQGFENRTIRLEAGIKPVYVFDGKPPDMKQELARYSKRDATDGLTEAVEVD 120
DB 95 TFSHLMGMEYRTIRMKENGKIPYI-----KSGELAKSELRAEAKFQLOQAAGCA 145
QY 121 KDAIEKSKRTYKYPYRQHNEDCKRLRLRMGVYVEASEAECAALCIDNKAFAVASED 180
DB 146 EQVVEKFTKMLVKYTKQNDCKHLSLMDGIPYDAPIEASCAALVAKGVYPATBED 205
QY 181 MDLTFGAPRFLRLHMDSSKRIIPVMEFDVAKVLELELTMDQFIDLCILCGDCYDSIK 240
DB 206 MDCTFGSPVLMQHLTASS-----DYCKSIW 231
QY 241 GIGGOTALKLIRHSGISIESILENINKROYIPEMDPYQEARLTFKPPNT-LDPELXMT 299
DB 232 STGPRAYDLQKHRSIEIYQRLDPNKYPLPENRLHKEAQLFLEPEVLGPESVELKMS 291
QY 300 ABDEGLISFLVKDNGFNEDRVYKAIEKIKSANKKSSQGRLESFFKPTATTSAPLKRKE- 358
DB 292 EPNBEELVKFVCGEKQFSEERHSGVKRLRTEKG-GGEPPGPDWTP--GRASPTKRPOH 348
QY 359 TSDKTSKAANKTKRAG 375
DB 349 QPDQHTSRSSSESHXG 365

RESULT 11
AAW24216
ID AAW24216 standard; Protein: 340 AA.
XX
XX AAW24216;
XX
XX 14-APR-1998 (first entry)
XX
XX Pyrococcus furiosus FEN-1 endonuclease.
XX
XX Nucleic acid cleavage: 5' nuclease; DNA cleavage; RNA cleavage;
XX invader directed cleavage; FEN-1; endonuclease.
XX
XX Pyrococcus furiosus.
XX
XX WO9727214-A1.
XX
XX 31-JUL-1997.
XX
XX 22-JAN-1997; 97WO-US01072.
XX
XX 02-DEC-1996; 96US-0759038.
XX 24-JAN-1996; 96US-0599491.
XX 12-JUL-1996; 96US-0682853.
XX 29-NOV-1996; 96US-0756386.
XX 02-DEC-1996; 96US-0758314.
XX
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichiev VI;
XX Olive DM, Prudent JR;
XX
XX WPI: 1997-393613/36.
XX N-PSDB: AAT76685.
XX
XX Thermostable structure-specific nuclease(s) - used for detection and
XX characterisation of nucleic acid sequences and variations in nucleic
XX acid sequences
XX
XX Example 28: Page 283-285; 457pp; English.
XX
XX This sequence comprises Pyrococcus furiosus (Pfu) FEN-1
XX endonuclease. Large-scale production of the enzyme was performed
```

using *E. coli* host cells transformed with a vector carrying the *Pfu* FEN-1 coding sequence (see AAT76685). *Pfu* FEN-1 is a thermostable CC enzyme. It can be used in novel methods for the detection and CC characterisation of nucleic acid sequences and variations in CC nucleic acid sequences.

Sequence 340 AA:

Query Match 33.4%; Score 648.5; DB 18; Length 340;  
Best Local Similarity 41.3%; Pred. No. 1.8e-51;  
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQFESYFGKRTAVDASMSIYQFLIVGRGMEITINAGEVYSHLOGMNRITRLLEA 78  
D 12 KEIENLYGKRTAIDALNAIYQFLSTIRQKDTPLMSKGRITSHLSGLFYRTINLMEA 71  
QY 79 GIKRVYVFDKPPDMKOEIAKRYSKRDATKDTLEAVGVGKDAIEKLSKRTVAVTROH 138  
D 72 GIKRVYVFDKPPDMKOEIAKRYSKRDATKDTLEAVGVGKDAIEKLSKRTVAVTROH 131  
QY 139 NEDCKRLRLMGVYVPAESEAECALCINDKVFVASEDMDSITFGAPRFLHMDP 198  
D 132 IEDAKKLELMGIPYQAPSEGEQAAYMAKGSYASASQDSDILFGAPRLVFNLTIT 191  
QY 199 SSKRIP-----VMEFDVAKVLELELTMDFIDICICGGDY-CDSTIKGIGG 244  
D 192 GKRLPGKNYVVEIKPELITILE---EVLEKLTIREKLEIALIVGTDYINPGIKIGL 247  
QY 245 QTALKLIRQHGSIESTILENLNKD--RYQIPEDWPYQEARLFKEPNVTLDIPELKWTAP 301  
D 248 KKALEIYRHH-----SKDPLAKFQKQSDVDLYAIKEFLNPVT-DNYNLWMDP 295  
QY 302 DEGLISFLKNGFNEDRYTKAIEKISAKNKSOGRLSEFPK 345  
D 296 DEGLISFLKNGFNEDRYTKAIEKISAKNKSOGRLSEFPK 339

RESULT 12  
AAW79970 ID AAW79970 standard; Protein; 340 AA.

XX AC AAW79970;  
XX DT 02-FEB-1999 (first entry)  
XX DE Pyrococcus furiosus FEN-1 endonuclease.  
XX KW Nucleic acid detection; multiple sequential invasive cleavage;  
XX FEN-1; endonuclease; nuclease.  
XX OS Pyrococcus furiosus.  
XX PN WO9842873-A1.  
XX PD 01-OCT-1998.  
XX PE 24-MAR-1998; 98WO-US05809.  
XX PR 24-MAR-1997; 97US-0823516.  
XX PA (THIR-) THIRD WAVE TECHNOLOGIES INC.  
XX PI Brow MAD, Hall JG, Kwiatkowski RM, Lyamlichev VI;  
XX PI Most AL, Vavra SH;  
XX DR WPI; 1998-557036/47.  
XX DR N-PSDB; AAV65840.  
XX PT Detecting target nucleic acid by sequence-specific cleavage of  
XX PT complex with two specific oligonucleotides - used to detect  
XX PT cytomagalovirus DNA  
XX PS Example 28b; Page 316-317; 524pp; English.

XX This is the amino acid sequence of FEN-1 endonuclease of CC  
XX Pyrococcus furiosus (Pfu). FEN-1 DNA (see AAV65840) has been CC  
XX ligated into vector for expression in *E. coli* cells. The invention CC  
XX relates to means for the detection and characterisation of nucleic CC  
XX acid sequences, and variations in nucleic acid sequences. It also CC  
XX relates to methods for forming a nucleic acid cleavage structure on CC  
XX a target sequence and cleaving this structure in a site-specific CC  
XX manner, preferably using a thermostable structure-specific nuclease CC  
XX such as FEN-1. cleavage of the cleavage structure by the CC  
XX nuclease indicates the presence of specific nucleic acid sequences CC  
XX or specific variants. The invention further relates to methods for CC  
XX the separation of nucleic acid molecules based on charge, methods CC  
XX for the detection of non-target cleavage products via the formation CC  
XX of a complete and activated protein binding region, and methods for CC  
XX the detection of nucleic acid from various viruses (e.g. human CC  
XX cytomagalovirus) in a sample. The method amplifies the detection CC  
XX molecule rather than the target itself, is less subject to CC  
XX contamination than exponential amplification processes, and allows CC  
XX many targets to be analysed in a single reaction.

Sequence 340 AA:

Query Match 33.4%; Score 648.5; DB 19; Length 340;  
Best Local Similarity 41.3%; Pred. No. 1.8e-51;  
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQFESYFGKRTAVDASMSIYQFLIVGRGMEITINAGEVYSHLOGMNRITRLLEA 78  
D 12 KEIENLYGKRTAIDALNAIYQFLSTIRQKDTPLMSKGRITSHLSGLFYRTINLMEA 71  
QY 79 GIKRVYVFDKPPDMKOEIAKRYSKRDATKDTLEAVGVGKDAIEKLSKRTVAVTROH 138  
D 72 GIKRVYVFDKPPDMKOEIAKRYSKRDATKDTLEAVGVGKDAIEKLSKRTVAVTROH 131  
QY 139 NEDCKRLRLMGVYVPAESEAECALCINDKVFVASEDMDSITFGAPRFLHMDP 198  
D 132 IEDAKKLELMGIPYQAPSEGEQAAYMAKGSYASASQDSDILFGAPRLVFNLTIT 191  
QY 199 SSKRIP-----VMEFDVAKVLELELTMDFIDICICGGDY-CDSTIKGIGG 244  
D 192 GKRLPGKNYVVEIKPELITILE---EVLEKLTIREKLEIALIVGTDYINPGIKIGL 247  
QY 245 QTALKLIRQHGSIESTILENLNKD--RYQIPEDWPYQEARLFKEPNVTLDIPELKWTAP 301  
D 248 KKALEIYRHH-----SKDPLAKFQKQSDVDLYAIKEFLNPVT-DNYNLWMDP 295  
QY 302 DEGLISFLKNGFNEDRYTKAIEKISAKNKSOGRLSEFPK 345  
D 296 DEGLISFLKNGFNEDRYTKAIEKISAKNKSOGRLSEFPK 339

RESULT 13  
AAW59940 ID AAW59940 standard; Protein; 340 AA.

XX AC AAW59940;  
XX DT 21-DEC-1998 (first entry)  
XX DE Amino acid sequence of the *Pfu* FEN-1 endonuclease.  
XX KW *Mja* FEN-1 endonuclease; Tag gene; structure-specific nuclease;  
XX KW mutant; DNA polymerase; bacteria; fungi; protozoa; RNA virus;  
XX KW hepatitis C virus; HCV; thermostable.  
XX OS Pyrococcus furiosus.  
XX PN WO98233774-A1.  
XX PD 04-JUN-1998.  
XX PR 26-NOV-1997; 97WO-US21783.



```

XX 02-DEC-1996: 96US-0758314.
PR 29-NOV-1996: 96US-0757653.
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX Kaiser MW, Lyamichev VI, Lyamicheva N;
PI WPI: 1998-322748/28.
DR N-PSDB: AAV53951.
XX Thermostable structure-specific nuclease(s) derived from mutant DNA
PT polymerase(s) - useful for detecting mutant allele(s) or strains of
PT microorganisms
XX Example 29: Pages 280-281; 472pp; English.
XX This is the amino acid sequence encoding the Pfu FEN-1 endonuclease,
CC used in the method of the invention. In this process thermostable
CC structure-specific nucleases are derived from mutant DNA polymerases,
CC which can be used for detecting mutant alleles or strains of
CC microorganisms. The structure-specific nucleases can be used in
CC mixtures, compositions and kits to treat nucleic acid, e.g. for
CC detection of wild type and mutant alleles of genes, for detection
CC and/or identification of strains of microorganisms such as bacteria,
CC fungi, protozoa, especially for detection of RNA viruses such as the
CC hepatitis C virus (HCV).
XX Sequence 340 AA:
SQ
Query Match 33.4%; Score 648.5; DB 19; Length 340;
Best Local Similarity 41.3%; Pred. No. 1.8e-51;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;
QY 19 KEQKESYFGKRIADVADSMSTIYOFILVGRGMEITLTNEAGEVTSHQMFNRTIRLEA 78
DB 12 KEIELNLYGKKIADIALNALIYOFILSTRQKDGTPLMDSKGRITSHLSGLFRTINLMEA 71
QY 79 GIKPYVFDGKPPDMKKQOELAKRYSKRDATKDLTEAVEVGDKDAIEKLSKRTVYTRQH 138
DB 72 GIKPYVFDGPPPEPKKELERREAREEAEKREAEKGEIEBARKYAOARATVEMNL 131
QY 139 NEDCKRLRLMGVVPVVEAPSEAECAALCINDKVFVAVASEMDSLTFGAPFLRLHMDP 198
DB 132 IEDAKKLELMGIPYVQASEGFAQAYMAKGSYASASQDYDSLTCGAPRLVNLTT 191
QY 199 SSKRIP-----VMEFVAKVLEDELTMDOFTIDLCICGCDY-CDSTKGTGG 244
DB 192 GKRKLPGKKNVYEIKPELTILE---EVKKEIKITREKILIEALIVGTDYNGGKIGL 247
QY 245 OTALKLIRQHGSIESTILENLNKD---RYQIPEDMPYQEARLRFKEPNVTLDIPELKTAP 301
DB 248 KKALEIVRH-----SKDPLAKFQKOSVDYLAIKEFLNLPVY-DNYNLVWMDP 295
QY 302 DEEGILSPFYKDNCFNEDRYTAKETIKSAKNNKSOGRLSEFPK 345
DB 296 DEEGILFLCDEHDSERKANKGLERLKAIKSGKOSTLESWPK 339
RESULT 14
AAW59953
ID AAW59953 standard; Protein: 326 AA.
XX
AC AAW59953;
XX
DT 22-DEC-1998 (first entry)
XX Amino acid sequence of the structure specific nuclease 5.
XX Tag mutant gene; thermostable; nuclease; mutant; DNA polymerase;
KW bacteria; protozoa; fungi; RNA virus; hepatitis C virus; HCV.
XX Synthetic.
OS

```

```

XX W09823774-A1.
XX 04-JUN-1998.
XX 26-NOV-1997: 97MO-US21783.
XX 02-DEC-1996: 96US-0758314.
PR 29-NOV-1996: 96US-0757653.
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX Kaiser MW, Lyamichev VI, Lyamicheva N;
PI WPI: 1998-322748/28.
XX Thermostable structure-specific nuclease(s) derived from mutant DNA
PT polymerase(s) - useful for detecting mutant allele(s) or strains of
PT microorganisms
XX Claim 1: Pages 331-332; 472pp; English.
XX This is the amino acid sequence of a structure specific nuclease
CC used in the method of the invention. In this process thermostable
CC structure-specific nucleases are derived from mutant DNA polymerases,
CC which can be used for detecting mutant alleles or strains of
CC microorganisms. The structure-specific nucleases can be used in
CC mixtures, compositions and kits to treat nucleic acid, e.g. for
CC detection of wild type and mutant alleles of genes, for detection
CC and/or identification of strains of microorganisms such as bacteria,
CC fungi, protozoa, especially for detection of RNA viruses such as the
CC hepatitis C virus (HCV).
XX Sequence 326 AA:
SQ
Query Match 33.4%; Score 647; DB 19; Length 326;
Best Local Similarity 42.6%; Pred. No. 2.4e-51;
Matches 140; Conservative 70; Mismatches 103; Indels 16; Gaps 7;
QY 19 KEQKESYFGKRIADVADSMSTIYOFILVGRGMEITLTNEAGEVTSHQMFNRTIRLEA 78
DB 12 KEIELNLYGKKIADIALNALIYOFILSTRQKDGTPLMDSKGRITSHLSGLFRTINLMEA 71
QY 79 GIKPYVFDGKPPDMKKQOELAKRYSKRDATKDLTEAVEVGDKDAIEKLSKRTVYTRQH 138
DB 72 GIKPYVFDGPPPEPKKELERREAREEAEKREAEKGEIEBARKYAOARATVEMNL 131
QY 139 NEDCKRLRLMGVVPVVEAPSEAECAALCINDKVFVAVASEMDSLTFGAPFLRLHMDP 198
DB 132 IEDAKKLELMGIPYVQASEGFAQAYMAKGSYASASQDYDALYLGAPRVYRNL-T 189
QY 199 SSKRIPVMEFVAKVLEDELTMDOFTIDLCICGCDY-CDSTKGTGGOTALKIRQHGSI 257
DB 190 TTKEMPEL-TEINVEFLDARISLDLIDLAIFMGIDYNGVKGKIGKRAYELVRS-GVA 247
QY 258 ESILENLKDRYQIPEDMPYQEARLRFKEPNVTLDIPELKTAPDEGLISFLYKONGEN 317
DB 248 KDVLK-----EVEYYDEIKRIKPEKVT-DNYSLSLSPDKGIGIKFLVDENDN 297
QY 318 EDRYTKAIEKIKS-AKNNKSOGRLSEFPK 345
DB 298 YDRVKRHVDKLYNLANKTKQKTLDAWFK 326
RESULT 15
AAW59951
ID AAW59951 standard; Protein: 332 AA.
XX
AC AAW59951;
XX
DT 22-DEC-1998 (first entry)
XX Amino acid sequence of the structure specific nuclease 3.
XX

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:27:44 ; Search time 16 seconds  
(without alignments)  
578.581 Million cell updates/sec

Title: US-09-805-311-2

Sequence: 1 MGKGLTKRLADNAPKAKE.....SDKTSKAANKTKAGGKKK 379

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: /cgn2\_6/prodata/1/1aa/5A.COMB.pep.\*
- 2: /cgn2\_6/prodata/1/1aa/5B.COMB.pep.\*
- 3: /cgn2\_6/prodata/1/1aa/6A.COMB.pep.\*
- 4: /cgn2\_6/prodata/1/1aa/6B.COMB.pep.\*
- 5: /cgn2\_6/prodata/1/1aa/PCTUS.COMB.pep.\*
- 6: /cgn2\_6/prodata/1/1aa/Backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1939	100.0	379	4	US-09-426-557-2	Sequence 2, Appl1
2	1939	100.0	379	4	US-09-426-557-6	Sequence 6, Appl1
3	1933	99.7	379	4	US-09-426-557-4	Sequence 4, Appl1
4	1933	99.7	379	4	US-09-426-557-8	Sequence 8, Appl1
5	1032	53.2	380	2	US-08-823-516-137	Sequence 137, App
6	1032	53.2	380	4	US-09-426-557-10	Sequence 10, Appl
7	1031	53.2	380	2	US-08-435-968E-1	Sequence 1, Appl
8	1014	52.3	378	2	US-08-823-516-138	Sequence 138, App
9	1007.5	52.0	377	2	US-08-455-968E-3	Sequence 3, Appl
10	996	51.4	378	2	US-08-455-968E-5	Sequence 5, Appl
11	938	48.4	382	2	US-08-823-516-119	Sequence 119, App
12	938	48.4	382	2	US-08-823-516-116	Sequence 116, App
13	648.5	33.4	340	2	US-08-823-516-79	Sequence 79, Appl
14	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
15	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
16	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
17	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
18	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
19	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
20	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
21	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
22	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
23	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
24	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
25	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
26	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App
27	648.5	33.4	340	2	US-08-823-516-116	Sequence 116, App

28	340	17.5	543	2	US-08-823-516-143	Sequence 143, App
29	334.5	17.3	530	2	US-08-823-516-142	Sequence 142, App
30	313.5	16.2	527	2	US-08-823-516-144	Sequence 144, App
31	205	10.6	872	1	US-08-766-014-2	Sequence 2, Appl1
32	184.5	9.5	291	4	US-09-105-697-6	Sequence 6, Appl1
33	173.5	8.9	834	5	PCT-US95-15327-2	Sequence 2, Appl1
34	173.5	8.9	548	2	US-08-484-956-86	Sequence 86, Appl
35	173	8.9	548	2	US-08-484-956-86	Sequence 86, Appl
36	173	8.9	695	2	US-08-484-956-87	Sequence 87, Appl
37	173	8.9	695	2	US-08-484-956-87	Sequence 87, Appl
38	173	8.9	810	4	US-09-587-856-2	Sequence 2, Appl1
39	173	8.9	810	4	US-09-587-856-2	Sequence 2, Appl1
40	173	8.9	810	4	US-09-587-856-2	Sequence 2, Appl1
41	173	8.9	810	4	US-09-587-856-2	Sequence 2, Appl1
42	173	8.9	832	1	US-07-977-538-2	Sequence 2, Appl1
43	173	8.9	832	1	US-07-977-538-2	Sequence 2, Appl1
44	173	8.9	832	1	US-08-156-020-2	Sequence 2, Appl1
45	173	8.9	832	1	US-08-156-020-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1  
US-09-426-557-2  
Sequence 2, Application US/09426557  
Patent No. 6232527  
GENERAL INFORMATION:  
APPLICANT: Mahajan, Pramod B.  
TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
FILE OF INVENTION: Theroof  
FILE REFERENCE: 0961  
CURRENT APPLICATION NUMBER: US/09/426,557  
CURRENT FILING DATE: 1999-10-22  
EARLIER APPLICATION NUMBER: 60/112,332  
EARLIER FILING DATE: 1998-12-15  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 379  
TYPE: PRT  
ORGANISM: Zea mays  
US-09-426-557-2

Query Match	Best Local Similarity	Score	1939;	DB 4;	Length	379;
Matches 379;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	MGKGLTKRLADNAPKAKEKEFESEYFGKRIAYDASMSIYOFELIVGRTGMETLNDAGE	60			
DB	1	MGKGLTKRLADNAPKAKEKEFESEYFGKRIAYDASMSIYOFELIVGRTGMETLNDAGE	60			
QY	61	VNSHLOGMNRITRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDDATYDLEAVEVGD	120			
DB	61	VNSHLOGMNRITRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDDATYDLEAVEVGD	120			
QY	121	KDAIEKLSKRYKVTROHNEDECKRLRLIMGVVVEAPEAECAALCINDKFAVAASED	180			
DB	121	KDAIEKLSKRYKVTROHNEDECKRLRLIMGVVVEAPEAECAALCINDKFAVAASED	180			
QY	181	MSDLTGARFRLHMDPSKKIPYMEFDVAKVLELELTMDOFTDLCILCGDYCDISIK	240			
DB	181	MSDLTGARFRLHMDPSKKIPYMEFDVAKVLELELTMDOFTDLCILCGDYCDISIK	240			
QY	241	GIGGOTALRLIHOGHSIESILENLKRDYQIPEDEMPQOEARLRFKEPNVTLDIPELKTZA	300			
DB	241	GIGGOTALRLIHOGHSIESILENLKRDYQIPEDEMPQOEARLRFKEPNVTLDIPELKTZA	300			
QY	301	PDEEGILSLVYDNGPNEDRYTKALEIKSAKNSSGORLESFFKPPATTSAPLKRKETS	360			
DB	301	PDEEGILSLVYDNGPNEDRYTKALEIKSAKNSSGORLESFFKPPATTSAPLKRKETS	360			
QY	361	DKTSKAANKTKAGGKKK 379				

Db 361 DKTSAANKKTAGGKK 379

## RESULT 2

US-09-426-557-6  
 ; Sequence 6, Application US/09426557  
 ; Patent No. 6232527  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mahajan, Pramod B.  
 ; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
 ; FILE REFERENCE: 0961  
 ; CURRENT APPLICATION NUMBER: US/09/426,557  
 ; EARLIER APPLICATION NUMBER: 60/112,332  
 ; EARLIER FILING DATE: 1998-12-15  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 6  
 ; LENGTH: 379  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 US-09-426-557-6

Query Match 100.0%; Score 1939; DB 4; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 1e-183;  
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKIAVDASMSIYQFLIVGRGTMETLTNEAGE 60  
 Db 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKIAVDASMSIYQFLIVGRGTMETLTNEAGE 60  
 Qy 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120  
 Db 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120  
 Qy 121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVYVVEAPSEAECAALCINDKVFVASED 180  
 Db 121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVYVVEAPSEAECAALCINDKVFVASED 180  
 Qy 181 MDSLTFGAPREFLRHLMDDPSKKIPVMEFDVAKVLELELTMDQFLDLCILGCDYCSIK 240  
 Db 181 MDSLTFGAPREFLRHLMDDPSKKIPVMEFDVAKVLELELTMDQFLDLCILGCDYCSIK 240  
 Qy 241 GIGGQFALKLIRHOGSIESILENLKDRYQIPEDWPYQEARRLKPEPNVTLDIPELKMTA 300  
 Db 241 GIGGQFALKLIRHOGSIESILENLKDRYQIPEDWPYQEARRLKPEPNVTLDIPELKMTA 300  
 Qy 301 PDEGLISFLVKNDFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
 Db 301 PDEGLISFLVKNDFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
 Qy 361 DKTSAANKKTAGGKK 379  
 Db 361 DKTSAANKKTAGGKK 379

## RESULT 3

US-09-426-557-4  
 ; Sequence 4, Application US/09426557  
 ; Patent No. 6232527  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mahajan, Pramod B.  
 ; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
 ; FILE REFERENCE: 0961  
 ; CURRENT APPLICATION NUMBER: US/09/426,557  
 ; EARLIER APPLICATION NUMBER: 60/112,332  
 ; EARLIER FILING DATE: 1998-12-15  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 4  
 ; LENGTH: 379  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 US-09-426-557-4

Query Match 99.7%; Score 1933; DB 4; Length 379;  
 Best Local Similarity 99.7%; Pred. No. 4e-183;  
 Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKIAVDASMSIYQFLIVGRGTMETLTNEAGE 60  
 Db 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKIAVDASMSIYQFLIVGRGTMETLTNEAGE 60  
 Qy 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120  
 Db 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120  
 Qy 121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVYVVEAPSEAECAALCINDKVFVASED 180  
 Db 121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVYVVEAPSEAECAALCINDKVFVASED 180  
 Qy 181 MDSLTFGAPREFLRHLMDDPSKKIPVMEFDVAKVLELELTMDQFLDLCILGCDYCSIK 240  
 Db 181 MDSLTFGAPREFLRHLMDDPSKKIPVMEFDVAKVLELELTMDQFLDLCILGCDYCSIK 240  
 Qy 241 GIGGQFALKLIRHOGSIESILENLKDRYQIPEDWPYQEARRLKPEPNVTLDIPELKMTA 300  
 Db 241 GIGGQFALKLIRHOGSIESILENLKDRYQIPEDWPYQEARRLKPEPNVTLDIPELKMTA 300  
 Qy 301 PDEGLISFLVKNDFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
 Db 301 PDEGLISFLVKNDFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
 Qy 361 DKTSAANKKTAGGKK 379  
 Db 361 DKTSAANKKTAGGKK 379

## RESULT 4

US-09-426-557-8  
 ; Sequence 8, Application US/09426557  
 ; Patent No. 6232527  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mahajan, Pramod B.  
 ; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
 ; FILE REFERENCE: 0961  
 ; CURRENT APPLICATION NUMBER: US/09/426,557  
 ; EARLIER APPLICATION NUMBER: 60/112,332  
 ; EARLIER FILING DATE: 1998-12-15  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 8  
 ; LENGTH: 379  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 US-09-426-557-8

Query Match 99.7%; Score 1933; DB 4; Length 379;  
 Best Local Similarity 99.7%; Pred. No. 4e-183;  
 Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKIAVDASMSIYQFLIVGRGTMETLTNEAGE 60  
 Db 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKIAVDASMSIYQFLIVGRGTMETLTNEAGE 60  
 Qy 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120  
 Db 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120  
 Qy 121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVYVVEAPSEAECAALCINDKVFVASED 180

```

|||||
Db 121 KDAIEKLSKRTVYKQHNEDCKRLRLMGVYVPEAPSEAEACALCINDKVFAYASED 180
QY 181 MDSLTFGAPRFLRLHMDSSKKIPVMEFDVAKYLELELTMDOFIDLCILCGDYCDISK 240
Db 181 KDSLTFGAPRFLRLHMDSSKKIPVMEFDVAKYLELELTMDOFIDLCILCGDYCDISK 240
QY 241 GIGGOTALKLIRHOGSISLLENLNKDRYOQIPEDMPYOEARRLFEKPNVTLIDIPELKMTA 300
Db 241 GIGGOTALKLIRHOGSISLLENLNKDRYOQIPEDMPYOEARRLFEKPNVTLIDIPELKMTA 300
QY 301 PDEELISFLVNDGNFNDRTYKALEIKSKAKNKSOGRLSEFPKPTATTSAPLKRRETS 360
Db 301 PDEELISFLVNDGNFNDRTYKALEIKSKAKNKSOGRLSEFPKPTATTSAPLKRRETS 360
QY 361 DKTSAANKKTKAGKKR 379
Db 361 DKTSAANKKTKAGKKR 379

RESULT 5
US-08-823-516-137
; Sequence 137, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FGRS-02736
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338

```

```

; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-823-516-137

Query Match 53.2%; Score 1032; DB 2; Length 380;
Best Local Similarity 53.8%; Pred. No. 6,1e-94;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGIGLTKRLADNAPKAMKEQKFESEYGRKRAVDASMSIYQFLIVYGRGMEITLNEAGE 60
Db 1 MGIGLTKRLADNAPSAIRENDIKSYGRKVAIDASMSIYQFLIAV-RQGGDVILQNEE 59
QY 61 VTSHLOGMFNTIRLLEAGTKPVYFPDGKPPDKKQFLARYSKRDDATKDLTEAVEVGD 120
Db 61 TTSHLMGFYRTIRNMEGIRPVYFDGKPPOLKSGELARSRERAEKQLOQAOAGA 119
QY 121 KDAIEKLSKRTVYKQHNEDCKRLRLMGVYVPEAPSEAEACALCINDKVFAYASED 180
Db 120 EOEVEKFTKRLVYKQHNEDCKRLSLMGIPYLDAPSEAEACALVYKAYATATED 179
QY 181 MDSLTFGAPRFLRLHMDSSKKIPVMEFDVAKYLELELTMDOFIDLCILCGDYCDISK 240
Db 180 MDCLTFGAPVLMRLHMTSEAKRLPIQELHLSRILOEGLNQEOPVDFICILGSDYCDISK 239
QY 241 GIGGOTALKLIRHOGSISLLENLNKDRYOQIPEDMPYOEARRLFEKPNVTLIDIPELKMTA 298
Db 240 GIGPRAVDLIQKHSIEIYVRIDPNKYPVPEWMLRKEAHQFLPEVY-LDEESVELKW 298
QY 299 TAPDEGLISFLVNDGNFNDRTYKALEIKSKAKNKSOGRLSEFPKPTATTSAPLKRRE 358
Db 299 SEPNEELIKFWGCKQPSERIRNSGVKRLSKSRQSGRLDDFRKVTGSLSS-AKRKE 357
QY 359 TSDK--TSKAANKTKYAGKKR 379
Db 358 PEPKSTKKKA---KTGAAGKFK 377

RESULT 6
US-09-426-557-10
; Sequence 10, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-426-557-10

Query Match 53.2%; Score 1032; DB 4; Length 380;
Best Local Similarity 53.8%; Pred. No. 6,1e-94;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGIGLTKRLADNAPKAMKEQKFESEYGRKRAVDASMSIYQFLIVYGRGMEITLNEAGE 60
Db 1 MGIGLTKRLADNAPSAIRENDIKSYGRKVAIDASMSIYQFLIAV-RQGGDVILQNEE 59
QY 61 VTSHLOGMFNTIRLLEAGTKPVYFPDGKPPDKKQFLARYSKRDDATKDLTEAVEVGD 120
Db 61 TTSHLMGFYRTIRNMEGIRPVYFDGKPPOLKSGELARSRERAEKQLOQAOAGA 119

```

Db 60 TTSHLGMENFRTIRRMENGIKPYVYFDGKRPOLKSGELAKRSERAEAKOLOQAAGA 119  
QY 121 KDAIEKLSKRYVKTROHNEDECKRLRLMGVAVVPAEAPSAEBCALCINDKVFVASED 180  
Db 120 EQEKEFTKLVKVTYOHNEDECKHLSMGITPYLDAPSAEASCALVKAAGVAAATED 179  
QY 181 MDSLTGAPRFLRHLMDPSSKRIPVNEFDVAKVLELELTMDQFDLCILGCGYCSIR 240  
Db 180 MDCLTFGSPVLMRHLTASAEAKKLPIQEFHLSRLQELGNGEQFVDCILGSDYCSIR 239  
QY 241 GIGGQATALKLIRHOGSIESTELENKDRYQIPEDMVPOEARLFEKPVYTD--IPELKW 298  
Db 240 GIGPKRAVDLIQHKRSIEELVRLDPNKIPVENMLHKAHOLELEPEV-LDPESVELKW 298  
QY 299 TAPDEGLISFLVKNGFNEDRVTKAIEKISAKNKSOGRLSEFPKPTATTSAPLKRE 358  
Db 299 SEPNEELIKFMGEGKOPSEERIRSGVKRLSKROGSTOGRLDPFKVTGSLSS-AKRE 357  
QY 359 TSDK--TSKAANKTKRAGKK 379  
Db 358 PEKGSTKKKA--KTGAAGKK 377

RESULT 7  
US-08-455-968E-1  
; Sequence 1, Application US/08455968E  
; Patent No. 5874283  
; GENERAL INFORMATION:  
; APPLICANT: Harrington, John L.  
; APPLICANT: Hsieh, Chih-Lin  
; APPLICANT: Lieber, Michael  
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/455,968E  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 18985-000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 380 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-455-968E-1

Query Match 53.2%; Score 1031; DB 2; Length 380;  
Best Local Similarity 53.8%; Pred. No. 7.7e-94;  
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;  
QY 1 MGIKGLTKLADNAPKAMEKQEFSEYFGKRIKIVDASMSIYQELIYVAGTETLTNAGE 60  
Db 1 MGIGLAKLADNAPSAIKRENDIKSFGRKVAIDASMSIYQELIYV-RGGGVTLQNEGE 59

QY 61 VTSHLOGMENFRTIRLEAGIKPYVYFDGKPPDMKKOELAKRSRRDDATKDLTEAVEGD 120  
Db 60 TTSHLGMENFRTIRRMENGIKPYVYFDGKRPOLKSGELAKRSERAEAKOLOQAAGA 119  
QY 121 KDAIEKLSKRYVKTROHNEDECKRLRLMGVAVVPAEAPSAEBCALCINDKVFVASED 180  
Db 120 EQEKEFTKLVKVTYOHNEDECKHLSMGITPYLDAPSAEASCALVKAAGVAAATED 179  
QY 181 MDSLTGAPRFLRHLMDPSSKRIPVNEFDVAKVLELELTMDQFDLCILGCGYCSIR 240  
Db 180 MDCLTFGSPVLMRHLTASAEAKKLPIQEFHLSRLQELGNGEQFVDCILGSDYCSIR 239  
QY 241 GIGGQATALKLIRHOGSIESTELENKDRYQIPEDMVPOEARLFEKPVYTD--IPELKW 298  
Db 240 GIGPKRAVDLIQHKRSIEELVRLDPNKIPVENMLHKAHOLELEPEV-LDPESVELKW 298  
QY 299 TAPDEGLISFLVKNGFNEDRVTKAIEKISAKNKSOGRLSEFPKPTATTSAPLKRE 358  
Db 299 SEPNEELIKFMGEGKOPSEERIRSGVKRLSKROGSTOGRLDPFKVTGSLSS-AKRE 357  
QY 359 TSDK--TSKAANKTKRAGKK 379  
Db 358 PEKGSTKKKA--KTGAAGKK 377

RESULT 8  
US-08-823-516-138  
; Sequence 138, Application US/08823516  
; Patent No. 5994069  
; GENERAL INFORMATION:  
; APPLICANT: Hall, Jeff G.  
; APPLICANT: Lyamichew, Victor I.  
; APPLICANT: Mast, Andrea L.  
; APPLICANT: Brox, Mary Ann D.  
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple  
; NUMBER OF SEQUENCES: 163  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/823,516  
; FILING DATE: 24-MAR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/01072  
; FILING DATE: 21-JAN-1997  
; APPLICATION NUMBER: US 08/759,038  
; FILING DATE: 02-DEC-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/758,314  
; FILING DATE: 02-DEC-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/756,386  
; FILING DATE: 29-NOV-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/682,853  
; FILING DATE: 12-JUL-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/599,491  
; FILING DATE: 24-JAN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027  
REFERENCE/DOCKET NUMBER: FORS-02736  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 378 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-08-823-516-138

Query Match 52.3%; Score 1014; DB 2; Length 378;  
Best Local Similarity 53.2%; Pred. No. 3,7e-92;  
Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps 6;

QY 1 MGKGLTLLADNAPKAKKEQKESYFGKRTAVDASMSIYQFLIVGRTGMEITLNEAGE 60  
DB 1 MGHLGLAKLIADVAPSAIRENDIKSYFGKVAIDASMSIYQFLIAV-RQGSDVLONEEGE 59  
QY 61 VTSHOGMNFRTIRLEAGIKPVYVFDGKPPDMKROELAKRYSKRDATKDLTEAVEVGD 120  
DB 60 TTS-LMGMEYRTIR-MENGIRKPVYVFDGKPPQLKSGELAKRSERAEKQLOQAQAGM 117  
QY 121 KDAIEKLSKRYKYTRQHNEDCKRLRLMGVYVEAESEAEACALCINDKYFAVASED 180  
DB 118 EEEVEFTRLVKTQHNDECKHLISLGMIPYLDABSEAEASCAALAKAGKYAAATED 177  
QY 181 MSLTFGARFLRLHMDSSKKIPMEFDVAKVLELEIMDQFIDLCICGDCYDSDIR 240  
DB 178 MCLTFGSPVLMRHLTASEAKKLPIQEPHLSRVLOELGINDQGVDCILGSDYCESIR 237  
QY 241 GIGGQATALKLRQHSISILENKNKRYQIPEDMPYQEARLKEPNVT-LDIPELKMT 299  
DB 238 GIGARAVDLQKHRSIEIVRLDPSKYVPENWMLHKEAQQLLEPEVDPESVELKWS 297  
QY 300 APDEGLISFLVKDGFEDRYTKAIEKIKSAKNSQGRLESFKEPTATTSAPLKRKT 359  
DB 298 ENEBELVAFMGCKEQFSEIRSGVKRLSKRSQSTGRDLDFKVTGSLSS-AKKKEP 356  
QY 360 SDKTSKAANKTKAGGKK 379  
DB 357 E---PRGPAKKRKTGAGK 373

## RESULT 9

US-08-455-968E-3

Sequence 3, Application US/08455968E

Patent No. 5874283

GENERAL INFORMATION:

APPLICANT: Harrington, John L.

APPLICANT: Hsieh, Chih-lin

APPLICANT: Lieber, Michael

TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

FILING DATE: 30-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 18985-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0300  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 377 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-455-968E-3

Query Match 52.0%; Score 1007.5; DB 2; Length 377;  
Best Local Similarity 53.4%; Pred. No. 1,6e-91;  
Matches 203; Conservative 68; Mismatches 100; Indels 9; Gaps 7;

QY 1 MGKGLTLLADNAPKAKKEQKESYFGKRTAVDASMSIYQFLIVGRTGMEITLNEAGE 60  
DB 1 MGHLGLAKLIADVAPSAIRENDIKSYFGKVAIDASMSIYQFLIAV-RQGSDVLONEEGE 59  
QY 61 VTSHOGMNFRTIRLEAGIKPVYVFDGKPPDMKROELAKRYSKRDATKDLTEAVEVGD 120  
DB 60 TTS-LMGMEYRTIR-MENGIRKPVYVFDGKPPQLKSGELAKRSERAEKQLOQAQAGM 116  
QY 121 KDAIEKLSKRYKYTRQHNEDCKRLRLMGVYVEAESEAEACALCINDKYFAVASED 180  
DB 117 MEEVEFTRLVKTQHNDECKHLISLGMIPYLDABSEAEASCAALAKAGKYAAATED 176  
QY 181 MSLTFGARFLRLHMDSSKKIPMEFDVAKVLELEIMDQFIDLCICGDCYDSDIR 240  
DB 177 MCLTFGSPVLMRHLTASEAKKLPIQEPHLSRVLOELGINDQGVDCILGSDYCESIR 236  
QY 241 GIGGQATALKLRQHSISILENKNKRYQIPEDMPYQEARLKEPNVT-LDIPELKMT 299  
DB 237 GIGARAVDLQKHRSIEIVRLDPSKYVPENWMLHKEAQQLLEPEVDPESVELKWS 296  
QY 300 APDEGLISFLVKDGFEDRYTKAIEKIKSAKNSQGRLESFKEPTATTSAPLKRKT 359  
DB 297 ENEBELVAFMGCKEQFSEIRSGVKRLSKRSQSTGRDLDFKVTGSLSS-AKKKEP 355  
QY 360 SDKTSKAANKTKAGGKK 379  
DB 356 EPKGS---AKKRAKTGAGK 372

## RESULT 10

US-08-455-968E-10

Sequence 10, Application US/08455968E

Patent No. 5874283

GENERAL INFORMATION:

APPLICANT: Harrington, John L.

APPLICANT: Hsieh, Chih-lin

TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/455,968E

```

; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0300
; TELEFAX: 415-576-0200
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-455-968E-10

Query Match      51.4%; Score 996; DB 2; Length 378;
Best Local Similarity 52.4%; Pred. No. 2.2e-90;
Matches 199; Conservative 70; Mismatches 103; Indels 8; Gaps 6;

1 MGIGLTKLADNAPKAMKEQKFESEYFGRIAVDASMSIYQFLIVYGRGMEITLNEAGE 60
1 MEHGLAKLIADYAPSAIRENDIKSFGRKVAIDASMSIYQFLIV-RGGGVDLQNEGE 59
61 VTSHLQGMENRTILLEAGIKPVYVFDGKPPDMKQOELAKRISKDDATKDLTEAVEVGD 120
60 TTS-LMGMEYFRMR-MENIKIPVYVFDGKPPOLKSGELAKRSERAEAKOLOAOQAGM 117
121 KDALIELSKRTYKVTROHNEDECKRLRLMGVNVPEAPSEAEACALCINDVFAVASED 180
118 EEEVEFTKRLVYKOHNDCKHLSTLNGIFPLDAPSAEASCAALAKAGVYAAATED 177
181 MSTLFGAPRFLRLHMDPSSKIPVMEFDYAVLELELTMDQIDLCILGCDYCSIR 240
178 MDLTCTGSPVLMRLHLSAFAKRLPIQEFHLISRYLOELGINOQFVDCITILGSDYCESIR 237
241 GIGGOTALKLIRHGSISTEILENDRQIPEDMFYQEARLFEKPNVNT-LDIEELKWT 299
238 GIGAKRVRDLQHKHSTIEELVRLDSKTPVPMNLHKEAQLLEPEVYADESVELKKS 297
300 APDEGLISFLVNDGNGFNEDRYTKALEIKSAKNSSGRLSEFFKPTATSAPLRKET 359
298 EPNEBELVYKMGCEKQFEFERIRSGVKRLSKSGSTQGRIDDFEYVTSLSLSS-AKRKEP 356
360 SDKTSRAANKTKRAGCKK 379
357 E---PKGPARKKATGTGAGK 373

RESULT 11
US-08-455-968E-5
; Sequence 5, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0300
; TELEFAX: 415-576-0200
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-455-968E-5

Query Match      48.4%; Score 938; DB 2; Length 382;
Best Local Similarity 49.5%; Pred. No. 1.2e-84;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

1 MGIGLTKLADNAPKAMKEQKFESEYFGRIAVDASMSIYQFLIVYGRGMEITLNEAGE 60
1 MGIGLNAIIEHVPASAIRKSDIKSFGRKVAIDASMSIYQFLIVARQODGQGLTNEAGE 60
61 VTSHLQGMENRTILLEAGIKPVYVFDGKPPDMKQOELAKRYSKRDDATKDLTEAVEVGD 120
61 TTSHLQGMENRTILLEAGIKPVYVFDGKPPDMKQOELAKRYSKRDDATKDLTEAVEVGD 120
121 KDALIELSKRTYKVTROHNEDECKRLRLMGVNVPEAPSEAEACALCINDKVFVASED 180
121 K---MKQERRLVYKSKHNEBAQKLTGLMGIPYLIAPTFAEACAEELAKGKYAAASED 177
181 MSTLFGAPRFLRLHMDPSSKIPVMEFDYAVLELELTMDQIDLCILGCDYCSIR 240
178 MDLTCTGSPVLMRLHLSAFAKRLPIQEFHLISRYLOELGINOQFVDCITILGSDYCESIR 237
241 GIGGOTALKLIRHGSISTEILENDRQIPEDMFYQEARLFEKPNVNTLDIPE 295
238 GIGAKRVRDLQHKHSTIEELVRLDSKTPVPMNLHKEAQLLEPEVYADESVELKKS 297
300 APDEGLISFLVNDGNGFNEDRYTKALEIKSAKNSSGRLSEFFKPTATSAPLRKET 359
298 EPNEBELVYKMGCEKQFEFERIRSGVKRLSKSGSTQGRIDDFEYVTSLSLSS-AKRKEP 356
360 SDKTSRAANKTKRAGCKK 379
357 E---PKGPARKKATGTGAGK 373

RESULT 12
US-08-823-516-139
; Sequence 139, Application US/08823516
; Patent No. 3994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichay, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
; TITLE OF INVENTION: Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medgen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/823,516  
FILING DATE: 24-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/01072  
FILING DATE: 21-JAN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/759,038  
FILING DATE: 02-DEC-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/758,314  
FILING DATE: 02-DEC-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/756,386  
FILING DATE: 29-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/682,853  
FILING DATE: 12-JUL-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/599,491  
FILING DATE: 24-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Ingolia, Diane E.  
REGISTRATION NUMBER: 40,027  
REFERENCE/DOCKET NUMBER: FORS-02736  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 139:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 382 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-08-823-516-139

Query Match 48.4%; Score 938; DB 2; Length 382;  
Best Local Similarity 49.5%; Pred. No. 1,2e-84;  
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

QY 1 MGIGLAKRLADNAPKMKQKQKESYFGKRIADVASMSTIYQFLIYVGRGTMETLTNAGE 60  
DB 1 MGIGLAAITSEHPSPAIRKSDIKSEFGKRVADASMSTLYDFLIARODGGLTNEGE 60  
QY 61 VTSHLQGMFNRTIRLLEAGIKPYVFDGKPPDMKQKOLAKRYSKRDDATKDLTEAVEYGD 120  
DB 61 TTSHLQGMFNRTIRLLEAGIKPYVFDGKPPDMKQKOLAKRYSKRDDATKDLTEAVEYGD 120  
QY 121 KDALEKLSKRTYVTRQHNEDCKRLRLMGVPEVFAEPAECALCTINDKVFVASED 180  
DB 121 K---MKERRLVYKSKHNEAOKLGLMGIPYITAPTEAKQCELAKKGVYAAASD 177  
QY 181 MDSTTFAPRRLHMDPSKKTPEWMEFYAKYLELELTMDQFDLCTLCCGCDYDSIK 240  
DB 178 MDVTCYTPPLRLHRTFSEKKEPIHEIDTELVRLGLDITIOFVLDLCTMCCDCESIR 237  
QY 241 GIGGOTAKLTIROHGSIESILENL-----NKDRYQIPEDMPYQEARLTFKEPNVTLDIPE 295  
DB 248 GVGVTALKLITKHGSIETLVEFTEGESNNTRKWIPEWPKYKQARMLFLDDEV-IDENE 296  
QY 296 --LKWTPADEGLISFLVKNQNGNEDRVTKAIEKISANNKSSQGRLESFFFTATT--- 350  
DB 297 INLWSPPKKELELYLDCDKTSEERKVSGLSLKKGKLSGLIGLDGFOVPVKTKEQ 356  
QY 351 -SAPLKRKETSDDTKSKAANKKTK 373  
DB 357 LAAAKRAQEKRLNK-NKMKVTK 379

RESULT 13  
US-08-757-653-176  
Sequence 176, Application US/08757653  
Patent No. 5843669  
GENERAL INFORMATION:  
APPLICANT: Kaiser, Michael W.  
APPLICANT: Lyamichev, Victor I.  
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using  
NUMBER OF SEQUENCES: 190  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Medien & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States Of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/757,653  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ingolia, Diane E.  
REGISTRATION NUMBER: 40,027  
REFERENCE/DOCKET NUMBER: FORS-02565  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 176:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-757-653-176

Query Match 33.4%; Score 648.5; DB 2; Length 340;  
Best Local Similarity 41.3%; Pred. No. 4.7e-56;  
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KQKQESYFGKRIADVASMSTIYQFLIYVGRGTMETLTNAGEVTSHLQGMFNRTIRLLEA 78  
DB 12 KEILENLGYKKRIADAIDANATIQEFLSTROKDGTPMDSKGRITSHLSGLFRTIIMLEA 71  
QY 79 GIKPYVFDGKPPDMKQKOLAKRYSKRDDATKDLTEAVEYGDGDAIEKLSKRTYVTRQH 138  
DB 72 GIKPYVFDGKPPDMKQKOLAKRYSKRDDATKDLTEAVEYGDGDAIEKLSKRTYVTRQH 131  
QY 139 NEDCKRLRLMGVPEVFAEPAECALCTINDKVFVASEDMDSLTFGAPRRLHMDP 198  
DB 132 IDEAKLLELMGIPVQAPSEGEAOAAVMAKGSVYASQDDVDSLFGAPRRLHMDP 191  
QY 199 SSKKIP-----VMEFDYAKYLELELTMDQFDLCTLCCGCD-CGSIIGG 244  
DB 192 GKRLLDGKRWYVEIKPELLILE---EVLKELKLTREKLTLELATIVGTQVNPFGIGIGL 247  
QY 245 GVALAKLIRHGSIESILENLKD---RYQIPEDMPYQEARLTFKEPNVTLDIPELKWTP 301  
DB 248 KALEIVRH-----SKDPLAKFQKOSQVDLYAIRKFEFLNPPVT-DNRYNLWRDP 295  
QY 302 DEGLISFLVKNQNGNEDRVTKAIEKISANNKSSQGRLESFFK 345  
DB 296 DEGLISFLVKNQNGNEDRVTKAIEKISANNKSSQGRLESFFK 339

RESULT 14

US-08-823-516-79  
: Sequence 79, Application US/08823516  
: Patent No. 5994069  
: GENERAL INFORMATION:  
: APPLICANT: Hall, Jeff G.  
: APPLICANT: Lyamichew, Victor I.  
: APPLICANT: Mast, Andrea L.  
: APPLICANT: Brow, Mary Ann D.  
: TITLE OF INVENTION: Detection of Nucleic Acids By Multiple  
: NUMBER OF SEQUENCES: 163  
: CORRESPONDENCE ADDRESSES:  
: ADDRESSEE: Medlen & Carroll, LLP  
: STREET: 220 Montgomery Street, Suite 2200  
: CITY: San Francisco  
: STATE: California  
: COUNTRY: United States Of America  
: ZIP: 94104  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/823,516  
: FILING DATE: 24-MAR-1997  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: PCT/US97/01072  
: FILING DATE: 21-JAN-1997  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/759,038  
: FILING DATE: 02-DEC-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/758,314  
: FILING DATE: 02-DEC-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/756,386  
: FILING DATE: 29-NOV-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/682,853  
: FILING DATE: 12-JUL-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/599,491  
: FILING DATE: 24-JAN-1996  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Ingolia, Diane E.  
: REGISTRATION NUMBER: 40,027  
: REFERENCE/DOCKET NUMBER: FORS-02736  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (415) 705-8410  
: TELEFAX: (415) 397-8338  
: INFORMATION FOR SEQ ID NO: 79:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 340 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: US-08-823-516-79

Query Match 33.4%; Score 648.5; DB 2; Length 340;  
Best Local Similarity 41.3%; Pred. No. 4.7e-56;  
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQFESYFGRIATAVDASMSIYQFLIVGRTGMEITLNEAGEVTSHTQGMENRTIRLEA 78  
Db 12 KEIELELVYKTKAIDALNAIYQFLSTIQKQGPFLMSKSRITSHLSGIFRTINIMEA 71  
QY 79 GIKRVYVFDKPPDMKKQELAKRYKRDATADLTLEAVVDKDAIEKLSKRYVKTROH 138  
Db 72 GIKRVYVFDSEPEFKKREAREKREALEKEGIEFKARYQATRVNEMK 131  
QY 139 NEDCKRLRLMGVPVVEAPSEAEACALCINDKVFVAVASEMDSLTFGAPRFLRLMDP 198

Db 132 IEDAKKLEIMGPIVQAPSECEAOAYMAKGSVYASASQDYSLLFGAPRLVRLTTR 191  
QY 199 SSKRIP-----VMEFDVAKVLELELTMOQFDLCTLCGCDV-COSTIGIG 244  
Db 192 GKRRLEKQNVYVTKPELLILE-----EYKELKLTREKLELIALVGTIDYINPGIGIGL 247  
QY 245 QYALKLIRQHSIESILEMKNKD---RYQIPEDWYQDARRLKEPNVTLIDELKWTAP 301  
Db 248 KKALEIVRH-----SKDPLAKFOQSDVDLYAIKEFFLMPVYT-DNYNLWRDP 295  
QY 302 DEAGLSIFVKGNGENEDRVTKAIEKRSKAKNKSOGRLSEFEK 345  
Db 296 DEEGILKFLQDEHDFEERVKNGLERLKKALKSKQSTLSWFK 339

RESULT 15  
US-08-823-516-136  
: Sequence 136, Application US/08823516  
: Patent No. 5994069  
: GENERAL INFORMATION:  
: APPLICANT: Hall, Jeff G.  
: APPLICANT: Lyamichew, Victor I.  
: APPLICANT: Mast, Andrea L.  
: APPLICANT: Brow, Mary Ann D.  
: TITLE OF INVENTION: Detection of Nucleic Acids By Multiple  
: NUMBER OF SEQUENCES: 163  
: CORRESPONDENCE ADDRESSES:  
: ADDRESSEE: Medlen & Carroll, LLP  
: STREET: 220 Montgomery Street, Suite 2200  
: CITY: San Francisco  
: STATE: California  
: COUNTRY: United States Of America  
: ZIP: 94104  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/823,516  
: FILING DATE: 24-MAR-1997  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: PCT/US97/01072  
: FILING DATE: 21-JAN-1997  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/759,038  
: FILING DATE: 02-DEC-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/758,314  
: FILING DATE: 02-DEC-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/756,386  
: FILING DATE: 29-NOV-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/682,853  
: FILING DATE: 12-JUL-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/599,491  
: FILING DATE: 24-JAN-1996  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Ingolia, Diane E.  
: REGISTRATION NUMBER: 40,027  
: REFERENCE/DOCKET NUMBER: FORS-02736  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (415) 705-8410  
: TELEFAX: (415) 397-8338  
: INFORMATION FOR SEQ ID NO: 136:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 340 amino acids  
: TYPE: amino acid



















A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-327 <STO>

A:Cross-references: GB:AE004437; NID:g10580870; PIDN:AG19690.1; GSPDB:GN00138

C:Genetics:

A:Gene: rad2

C:Superfamily: conserved hypothetical protein YK113C

Query Match 19.0%; Score 369; DB 2; Length 327;

Best Local Similarity 31.8%; Pred. No. 2e-17;  
Matches 106; Conservative 60; Mismatches 139; Indels 28; Gaps 9;

```
Oy 17 AMKQKFESEYGRKIAVDASMSIYFL-IVGRTGMETLTNEAGEVTSILOGMFNRTIRL 75
Db 11 AIEETPFADLEGSVVAVDANMILYKYLTTVQMTGADVTTSDGTEVANLVGAVGCLPKF 70
Oy 76 LEAGIKRYVYFDGKFPDMKKQELAKRYSKRDATKDLTEAVEYGDKAIEKSKRTVAVT 135
Db 71 FEHGITPVFWDDGGVTCLKDDEIADREQREYEEQLDDAREGDAEAERLIDARTORLT 130
Oy 136 ROHNEDCKRLRLMGVPEVEAPEBAECAALC-INDKVFAVASEDMDSLTFGAPRFLRH 194
Db 131 PTIHETTFELFDLIDIPQVEAPAEGEQAAYMTRTDDAVDAGSDDYDCLLGSPVTLRO 190
Oy 195 LMDPSKKIPYMERPDVAKVLELELTMDQFIDLCILGCGDYCDISYKIGIGOTALKLIRH 254
Db 191 L--TSSGHPELMDPA--TLAEHDLWEQLVDVGLIGTDFNPGIDGFGFTTALDAIGBH 246
Oy 255 GSIESIL---ENL-NKDRYQIPEDWPEQEARRLFEKPEPVTLDI---PELKWTAPEDEGL 306
Db 247 GDLMDVYLAEGEHVAHGDR-----IRELFLNPVTDYVDPV---SPALDAA 292
Oy 307 ISFLVKDNGFNEEDRVTKAIEKISAKNKSQGR 339
Db 293 RAFVTDMEVNDADAVARGFERIDAAAQGTGLDR 325
```

Search completed: November 5, 2002, 16:29:44  
Job time : 24 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:25:09 : Search time 15 seconds  
(without alignments)  
978.314 Million cell updates/sec

Title: US-09-805-311-2

Sequence: 1 MGIGLTKRLADNAPKAKE.....SDKTSKAAKKTKAGKKK 379

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Swissprot\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1032	53.2	380	1	FEN1_HUMAN
2	1023.5	52.8	380	1	RAD2_SCHPO
3	1014	52.3	378	1	FEN1_MOUSE
4	998	48.4	382	1	RA27_YEAST
5	560.5	28.9	326	1	FEN1_METUA
6	260.5	13.4	571	1	EXO1_SCHPO
7	245.5	12.7	496	1	YAD3_SCHPO
8	232.5	12.0	1031	1	RAD2_YEAST
9	220.5	11.4	1186	1	XPG_HUMAN
10	220	11.3	1170	1	XPG_MOUSE
11	218	11.2	702	1	EXO1_YEAST
12	204.5	10.5	430	1	DIN7_YEAST
13	204.5	10.5	1112	1	RA13_SCHPO
14	200	10.3	1196	1	XPG_XENLA
15	184.5	9.5	834	1	DPOL_THECA
16	176.5	9.1	986	1	DPOL_SYNY3
17	173	8.9	832	1	DPOL_THEAO
18	172.5	8.9	834	1	DPOL_THERH
19	166	8.6	289	1	EX53_AQUAE
20	157.5	8.1	831	1	DPOL_THERFL
21	156.5	8.1	833	1	DPOL_THERI
22	154.5	8.0	877	1	DPOL_STRPN
23	153.5	7.9	877	1	DPOL_LACIC
24	151.5	7.8	759	1	YEN1_YEAST
25	149	7.7	930	1	DPOL_HAEIN
26	144	7.4	877	1	DPOL_BACCA
27	142.5	7.3	877	1	DPOL_LACIA
28	138	7.1	921	1	DPOL_RICHE
29	135.5	7.0	880	1	DPOL_BACSU
30	134.5	6.9	956	1	DPOL_DEIRA
31	134	6.9	867	1	DPOL_RICPR
32	134	6.9	922	1	DPOL_RICFE
33	123	6.3	876	1	DPOL_BACST

34	118.5	6.1	850	1	DPOL_ANATH	059156 anaerocellu
35	113.5	5.9	942	1	DPOL_CHLAU	008307 chloroflexu
36	112	5.8	317	1	ACCO_PEA	P31239 piscum saliv
37	112	5.8	984	1	SECA_AQUAE	067718 aquilex aeo
38	109.5	5.6	891	1	DPOL_HELPU	P56105 helicobacte
39	109.5	5.6	1016	1	DPOL_RHILE	Q98192 rhizobium 1
40	109.5	5.6	1650	1	BP28_CAEEL	Q23495 caenorhabdi
41	109	5.6	886	1	RA50_ARCFU	029230 archaeoglob
42	109	5.6	897	1	DPOL_HELPU	092139 helicobacte
43	108.5	5.6	444	1	TIG_CAMTE	Q46108 campylobact
44	108.5	5.6	928	1	DPOL_ECOLI	P00582 escherichia
45	108.5	5.6	1509	1	MKSN_ACRCA	P05659 acanthameb

## ALIGNMENTS

```

RESULT 1
ID FEN1_HUMAN STANDARD; PRT; 380 AA.
AC P39748;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE FIAP endonuclease-1 (Maturation factor 1) (MFI).
GN FEN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94277093; PubMed=8007985;
RA Murray J.M., Tavassoli M., Al-Harithy R., Sheldrick K.S.,
RA Lehman A.R., Carr A.M., Watts F.Z.;
RT "Structural and functional conservation of the human homolog of the
RT Schizosaccharomyces pombe rad2 gene, which is required for chromosome
RT segregation and recovery from DNA damage";
RL Mol. Cell. Biol. 14:4878-4888(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphocytes;
RA Lieber M.R., Hsieh C.L., Harrington J.J.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Coleman M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Gaines J.,
RA Dangnan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Altix C., Andrieux T., Frankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Krommiller B., Ariello A., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S.;
RA Carrano A.V.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=95050647; PubMed=7961795;
RA Robins P., Pappin D.J., Wood R.D., Lindahl T.;
RT "Structural and functional homology between mammalian DNase IV and
RT the 5'-nuclease domain of Escherichia coli DNA polymerase I.";
RL J. Biol. Chem. 269:28535-28538(1994).
CC -I- FUNCTION: ENDONUCLEASE THAT CLEAVE 5'FLAP STRUCTURE AND FAILS
CC TO CLEAVE OTHER DNA STRUCTURES, INCLUDING 3'FLAPS AND SINGLE
CC STRANDED DNA (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -I- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1

```

```

CC      SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: X76771; CA554166.1; -
DR      EMBL: I37374; AA91331.1; -
DR      EMBL: AC004770; AAC2394.1; -
DR      EMBL: BC000323; AAH00323.1; -
DR      HSSP: Q58839; 1A76.
DR      MIM: 600393; -
DR      InterPro: IPR002421; 5_3_exonuclease.
DR      InterPro: IPR000513; Exo_N.I.
DR      InterPro: IPR003584; HNH_2.
DR      InterPro: IPR001532; XPG_1.
DR      Pfam: PF01367; 5_3_exonuclease; 1.
DR      Pfam: PF00867; XPG_1; 1.
DR      Pfam: PF00752; XPG_N; 1.
DR      PRINTS: PR00853; XPGRADSUPER.
DR      SMART: SM00279; HNH2; 1.
DR      SMART: SM00484; XPG1; 1.
DR      SMART: SM00485; XPGN; 1.
DR      PROSITE: PS00841; XPG_1; 1.
DR      PROSITE: PS00842; XPG_2; 1.
DR      Hydrolase: Nuclease; Endonuclease; Nuclear protein.
FT      DOMAIN 1 104
FT      DOMAIN 122 253
SQ      SEQUENCE 380 AA; 42593 MW; 5154PF6E5792C5 CRC64;

Query Match      53.2%; Score 1032; DB 1; Length 380;
Best Local Similarity 53.8%; Pred. No. 3,1e-61;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY      1 MGIGKLTLLADNAPKAMKEQKFESEYFGKRIKAVDASMSIYQFLVVGRTGMETLTNEAGE 60
DB      1 MGIGGLAKLLIADVAPSAIRENDIKSYFGKRVADIASMSIYQFLIIV-RGGGVQLQNEGE 59

QY      61 VTSHLQGMFKRTIRLEAGIKRPVYVFDGKPRPMKKOELAKRYSKRDATKDLTEAVEGCD 120
DB      60 TTSHLQGMFKRTIRLMENGLIKRPVYVFDGKPRPOLKSGELAKRSEKRAEAKOLOQQAAGA 119

QY      121 KDAIEKLSKRTYVAVTROHNEDCKRLRLMGVYVEAPSEAEACAALCINDKVFVAVASD 180
DB      120 EQEVEKFTKRLVAVTROHNEDCKHLLSLMGIRPLDAPSEAEACAALVAKGVYVAATED 179

QY      181 MDSLTFGARPLRLHMDPSSKKIPVWEFDVAKVLELELTMQFDLDCILCCGYCDSIK 240
DB      180 MDSLTFGSPVLRHRLVLAASEKKLPVIOEFHLISRLDELGNQDFVLDCLILGSDDCESIR 239

QY      241 GIGGQATLKIIRQHSIESILENLKRDYOIPEDWYQEARLFEKPNVTTD--IPELKW 298
DB      240 GIGKRAVVDLQHKHSEELVRLDNPKNKRPVENNLAKRAHDLFLDEPV-LDPEVELEKW 298

QY      299 TAPDEGLIFLVKDNQFNDRVTKAIEKIKSAKKNSSQGRLESPFKPATTSAPLAKKE 358
DB      299 SEPHEELIFMCGEKQFSEERIRSGVKRLSSKROSTQGRLDPEFKYGLSS--AKRKE 357

QY      359 TSDK--TSKAANKTKAGGKK 379
DB      358 PERKGTAKKA---KTGAAGKFK 377

```

```

DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      DNA repair protein rad2.
GN      RAD2 OR SPAC36.06C.
OS      Schizosaccharomyces pombe (Fission yeast).
OC      Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC      Schizosaccharomycetales; Schizosaccharomycetaceae;
OC      Schizosaccharomycetes.
OX      NCBI_TaxID=4896;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94277093; PubMed=8007985;
RA      Murray J.M., Tavassoli M., Al-Harithy R., Sheldrick K.S.,
RA      Lehmann A.R., Carr A.M., Watts F.Z.;
RT      "Structural and functional conservation of the human homolog of the
RT      Schizosaccharomycetes pombe rad2 gene, which is required for chromosome
RT      segregation and recovery from DNA damage.";
RL      Mol. Cell. Biol. 14:4878-4888(1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=972;
RA      Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL      Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
CC      -!- FUNCTION: REQUIRED FOR THE FIDELITY OF CHROMOSOME SEPARATION AT
CC      MITOSIS AND WHICH IS ALSO INVOLVED IN THE RESPONSE TO DNA DAMAGE.
CC      -!- SUBCELLULAR LOCATION: Nuclear (probable).
CC      -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENONUCLEASE FAMILY. FEN1
CC      SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: X77041; CAB36991.1; -
DR      EMBL: Z99167; CAB16282.1; -
DR      HSSP: Q58839; 1A76.
DR      InterPro: IPR000513; Exo_N.1.
DR      InterPro: IPR003584; HNH_2.
DR      InterPro: IPR001532; XPG_1.
DR      Pfam: PF00867; XPG_1; 1.
DR      Pfam: PF00752; XPG_N; 1.
DR      PRINTS: PR00853; XPGRADSUPER.
DR      SMART: SM00279; HNH2; 1.
DR      SMART: SM00484; XPG1; 1.
DR      SMART: SM00485; XPGN; 1.
DR      PROSITE: PS00841; XPG_1; 1.
DR      PROSITE: PS00842; XPG_2; 1.
DR      DNA repair: Hydrolase; Nuclease; Endonuclease; Nuclear protein.
FT      DOMAIN 1 105
FT      DOMAIN 123 254
SQ      SEQUENCE 380 AA; 42866 MW; 36040659B934CBE CRC64;

Query Match      52.8%; Score 1023.5; DB 1; Length 380;
Best Local Similarity 50.7%; Pred. No. 1,1e-60;
Matches 194; Conservative 73; Mismatches 107; Indels 9; Gaps 3;

QY      1 MGIGKLTLLADNAPKAMKEQKFESEYFGKRIKAVDASMSIYQFLVVGRTGMETLTNEAGE 60
DB      1 MGIGKLAQVLSHAPASVKNHNDIKNFGRKVAIDASMSIYQFLIOVSODQOQLMNEGE 60

QY      61 VTSHLQGMFKRTIRLEAGIKRPVYVFDGKPRPMKKOELAKRYSKRDATKDLTEAVEGCD 120
DB      61 TTSHLQGMFKRTIRLVNDGKIPCFVFDGKPRPLKSGELAKRVAHQAKREDEETKEVGT 120

QY      121 KDAIEKLSKRTYVAVTROHNEDCKRLRLMGVYVEAPSEAEACAALCINDKVFVAVASD 180
DB      121 AAVYDFAKRTYVAVTROHNEDCKRLLELMGIPVNAPEAEACAALARSQVYVAASD 180

QY      181 MDSLTFGARPLRLHMDPSSKKIPVWEFDVAKVLELELTMQFDLDCILCCGYCDSIK 240
DB      181 MDSLTFGARPLRLHMDPSSKKIPVWEFDVAKVLELELTMQFDLDCILCCGYCDSIK 240

```



```

CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S93804; AAB21998.1; -
CC DR EMBL: Z28113; CA8193.1; -
CC DR PIR: S22267; S22267.
CC DR HSSP: O58839; 1A76.
CC DR SGD: S0001596; RAD27.
CC DR InterPro: IPR002421; 5_3_exonuclease.
CC DR InterPro: IPR000513; EXO_N.1.
CC DR InterPro: IPR003584; HHN_2.
CC DR InterPro: IPR001532; XPG_1.
CC DR Pfam: PF01367; 5_3_exonuclease; 1.
CC DR Pfam: PF00867; XPG_1; 1.
CC DR Pfam: PF00752; XPG_N; 1.
CC DR PRINTS: PRO0853; XPGRADSUPER.
CC DR SMART: SM00279; HHN2; 1.
CC DR SMART: SM00484; XPG1; 1.
CC DR SMART: SM00485; XPGN; 1.
CC DR PROSITE: PS00841; XPG_1; 1.
CC DR PROSITE: PS00842; XPG_2; 1.
CC DR DNA_repair: Hydrolase; Nuclease; Endonuclease; Nuclear protein.
CC FT DOMAIN 1 105 N-DOMAIN.
CC FT SEQUENCE 382 AA; 43279 MW; 1F54B08720121C8C CRC64;
CC SQ
Query Match 48.4%; Score 938; DB 1; Length 382;
Best Local Similarity 49.5%; Pred. No. 5e-55;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;
OY 1 MGIGLTLADNAPKAKKEQFESYGRKTAVDASITVQFLVGRGMEITLTFNAGE 60
DB 1 MGIGLNLITSEHVSATRKSDIKSFGRKVAIDASHSLVQFLAVNQDGGGLTNAGE 60
OY 61 VTSHLOGMFNTIRLLEAGIKPRVYFDGKPPDMKKOELAKRYKRDDATDLEAVEVG 120
DB 61 TTSHLMGMFYTLRIIDGICPCVYFDGKPPDLKSHLETRSSRVETKEKLAETTELE 120
OY 121 KDAIEKLSKRYVKTROINECKRLRLMGVYVABSEAEACALCINDKVAVAASED 180
DB 121 K---KKOERRRLVSKSEHNEEAOKLLGMIPIYIIAPEAEAOCAELAKKGVYAAASED 177
OY 121 K---KKOERRRLVSKSEHNEEAOKLLGMIPIYIIAPEAEAOCAELAKKGVYAAASED 177
OY 181 MSLTFGAPRFLRHLMDPSSKKRIPWEDVAKVLEELTMDQFIDLCILCCDVCDSIK 240
DB 178 MDTLCYRPPFLRLHLPFSEAKKEPHIEDTLEVLRLGDLITIEQFDLCIMGCGYCESTR 237
OY 241 GIGGOTATLKLROHGSIESILENT-----NKDRYQIPEDWPYQEARRLFEPNVTLDIPE 295
DB 238 GVGPTATLKLITKTHGSIEKIVFIESGESNNTKWKIPEDWPYQKARMLFDLPEV-IDENE 296
OY 296 --LAKTAPDEEGILSFVKNQNGFNEDRYTKATEIKSAKNNSSQGRLESPFKPATY--- 350
DB 297 INIKSPPEKEKLELYLDCDDKFESEKSGISRLKGLKSGIQGRIDGFFOVVPKTREQ 356
OY 351 -SAPLKRKETSDDTKSAANKKTK 373
DB 357 LAAAKRAQENKKLNK-NKKYTK 379

```

```

DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Flap endonuclease.
GN FEN OR M0144.
OS Methanococcus jannaschii.
CC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
CC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A. 2661 / ATCC 43067;
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merriack J.M., Nguyen D.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Huysl M.A., Kalne B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RT Science 273:1058-1073(1996).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=98440432; PubMed=9765572;
RA Rao H.S., Rosenfeld A., Wetmur J.G.;
RT "Methanococcus jannaschii flap endonuclease: expression, purification,
RT and substrate requirements";
RT J. Bacteriol. 180:5406-5412(1998).
RN [3]
RP CHARACTERIZATION. AND X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=99200693; PubMed=10102570;
RA Bae K.W., Bae K.W., Cho C.S., Hwang K.Y., Kim H.-R., Sung H.-C.,
RA Cho Y.;
RT "Expression, purification, characterization and crystallization of
RT flap endonuclease-1 from Methanococcus jannaschii.";
RT Mol. Cells 9:45-48(1999).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=98363215; PubMed=9699635;
RA Hwang K.Y., Bae K., Kim H.-Y., Cho Y.;
RT "The crystal structure of flap endonuclease-1 from Methanococcus
RT jannaschii.";
RT Nat. Struct. Biol. 5:707-713(1998).
RN [5]
RP FUNCTION: Has 5' endo-/exonuclease and 5' pseudo-Y-endonuclease
CC activities. Cleaves the junction between single and double-
CC stranded regions of flap DNA.
CC -1- COFACTOR: MAGNESIUM.
CC -1- MISCELLANEOUS: ITS OPTIMUM PH IS 6-7.
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67585; AAB99454.1; -
CC DR PDB: 1A76; 03-AUG-99.
CC DR PDB: 1A77; 03-AUG-99.
CC DR TIGR: M01444; -
CC DR InterPro: IPR000513; EXO_N.1.
CC DR InterPro: IPR003584; HHN_2.
CC DR InterPro: IPR001532; XPG_1.
CC DR Pfam: PF00867; XPG_1; 1.
CC DR Pfam: PF00752; XPG_N; 1.
CC DR PRINTS: PRO0853; XPGRADSUPER.
CC DR SMART: SM00279; HHN2; 1.
CC DR SMART: SM00484; XPG1; 1.
CC DR SMART: SM00485; XPGN; 1.

```



KM Hydrolase; Nuclease; Endonuclease; Magnesium; 3D-structure;  
 KW Complete proteome.  
 FT METAL 154 MAGNESIUM.  
 SQ SEQUENCE 326 AA; 37527 MW; 311427F2B4B67580 CRC64;

Query Match 28.9%; Score 560.5; DB 1; Length 326;  
 Best Local Similarity 39.1%; Pred. No. 3; se 30;  
 Matches 127; Conservative 66; Mismatches 114; Indels 16; Gaps 7;

QY 23 PESYFGKRIADVASMSIYQFLIVGRTGEMTLNAGEVTSHLQGMFNRTILLEGKIP 82  
 16 FEDLKGKKAIDQMNALQFLTSIRLDGSPLRNKGCEITSAYNGVYKTHLEMDINP 75  
 DB 83 VVFGKPPDMKQKQELARYSKRDDATKDTLEAVEVGDMAIKRLSKRPVYKTRQHNEDC 142  
 76 IVVFGGEPKLEKTRKRRKREKKAELKKEKDEFEAKYAKRSYILPKAVENC 135  
 QY 143 KRLRLMGVYVVEAPSEAEACALCINDKYFAVASDMDSLTFGAPRLRLHMDPSSK 202  
 136 KLLSLMGIPYVEAPSEAEACALCINDKYFAVASDMDSLTFGAPRLRLHMDPSSK 202  
 DB 203 IVMEEDYAKVLELELMDQFIDLCICGDY-CDSIKIGGQTLAKILRQHSIESIL 261  
 194 MEL-IELENEVEDLRISLDLIDIAIFMGTDYMGSGVKIGKRAVELYRS-GVAKDVL 251  
 QY 262 ENLKDNDROIPEDMYQARLRFKPNVTLDIPELKATPDEGLISFLYKDNQFENEDV 321  
 252 KR-----EVEYDIDIKRIFEKPKVT-DNYSLSLKLPDKELIKELVDENDENYDRV 301  
 DB 322 TKAIEKRS-AKKNSSQGRLESFPR 345  
 302 KKHVDKLYNLANKTKOKTLDAMFK 326

RESULT 6  
 EXOL\_SCHPO STANDARD: PRT: 571 AA.  
 ID EXOL\_SCHPO  
 AC P53695:  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Exonuclease I (EXO I).  
 GN EXOI OR SPIC29A10.05.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RX MEDLINE=95158876; PubMed=7855597;  
 RA Szankasi P., Smith G.R.;  
 RT "A role for exonuclease I from S. pombe in mutation avoidance and  
 RT mismatch correction";  
 RL Science 267:1166-1169(1995).  
 RN 12)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: 5'->3' DOUBLE-STRANDED DNA EXONUCLEASE THAT COULD ACT IN  
 CC A PATHWAY THAT CORRECTS MISMATCHED BASE PAIRS.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (potential).  
 CC -1- INDUCTION: BY MEIOSIS.  
 CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

DR EMBL: L35174; AAC1648.1; -  
 DR EMBL: AL034463; CA22433.1; -  
 DR InterPro: IPR000513; EXO\_N\_I.  
 DR InterPro: IPR001584; HHH\_2.  
 DR InterPro: IPR001440; TPR.  
 DR InterPro: IPR001532; XPG\_I.  
 DR Pfam: PF00867; XPG\_I.1.  
 DR Pfam: PF00752; XPG\_N.1.  
 DR PRINTS: PR00853; XPGRADSUPER.  
 DR SMART: SM00279; HH2.1.  
 DR SMART: SM00484; XPGI.1.  
 DR SMART: SM00485; XPGI.1.  
 DR PROSITE: PS00841; XPG\_1.1.  
 DR PROSITE: PS00842; XPG\_2.1.  
 KM DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.  
 FT DOMAIN 1  
 FT DOMAIN 1  
 FT SEQUENCE 571 AA; 63867 MW; 7D21B3778104BA9D CRC64;

Query Match 13.4%; Score 260.5; DB 1; Length 571;  
 Best Local Similarity 25.4%; Pred. No. 4e-10;  
 Matches 110; Conservative 73; Mismatches 167; Indels 83; Gaps 17;

QY 1 MGIRGLTLADNAPKAM-KEQKFESYFGKRIADVASMSIYQFLIVGRTGEMTLNEAG 59  
 1 MGIRGLTLGL---KPKQSSHVEEFSGKTLGVGVWLKRAVTCAN--ELAEFK-- 50  
 DB 60 EVYSHLQGMENRTIRLENGIRPVYVFDCKPPDMKQKQELARYSKRDDATKDTLEAVEVG 119  
 51 ETRKYLAETHOALMOYGVKPLVFDGRLPKCASTEQRKRRO-----EAEELG 103  
 QY 120 DK-----DALEKSRRTVAVTRQHNEDCKRLRLMGVYVVEAPSEAEACALCIND 171  
 104 KRLMDCKSKSQALMORS-KCVDTPEMAKRLILAREGISIYAPYBADNOVYLEKEN 162  
 QY 172 KVEAVASEMDSLTFGAPRLRLHMDPSSKTIPEVEPVAVLE-ELSLTMDQFIDLCIL 230  
 163 IIDGILTEBDSMDLVFCAQVVL-EKMDGFNCITIRRNDAVADLINRLPEKRLHMAIF 221  
 QY 231 CGCDYCSIKIGGQYALILRQHSIESILENLKDR-YOIPEDWYQEA-----R 281  
 222 SGCDYTDGAVGMKLTALVLTQKYPEPRAIRAMRLDKSLAVPSFEKFPALADLARHQ 281  
 DB 282 RLKEPEPNT---LDIPELMTAPDEGLISFLYKDNF-----NEDRYKALETKSA 331  
 282 RYVCPKDKTILVHLSPEERLSVHEDAFISFP--DNQALDAIDAGRSNPITTKCAFIDKDS 339  
 DB 332 KNS-----SQRLESFF-----KPTATTS-----APLKRKE 358  
 QY 340 SMQSFYKTTITTSKRGISKTDISNEFKSIIPSKRPTKSTSLDVTVNVKQRTHLANDI 399  
 QY 359 TSDXTSKAANKK 371  
 DB 400 SSKQSIKSNK 412

RESULT 7  
 Y31\_SCHPO STANDARD: PRT: 496 AA.  
 ID Y31\_SCHPO  
 AC 009708:  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical 57.8 kDa protein Clb11.01c in chromosome I.  
 GN SPAC18B11.01c.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-972:
RA Devlin K., Churche C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z50728; CAA90586.1; -
DR InterPro: IPR000513; Exo_N.1.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR001532; XPG_1.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XGRADSUPER.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
RW Hypothetical protein; Hydrolase; Nuclease; Endonuclease.
SQ SEQUENCE 496 AA; 57763 MW; F2A16D26846F24C3 CRC64;

Query Match 12.7%; Score 245.5; DB 1; Length 496;
Best Local Similarity 30.2%; Pred. No. 3,3e-09;
Matches 62; Conservative 47; Mismatches 77; Indels 19; Gaps 5;

OY 111 DLTAEVYGKDAIEKISRTAVYVTOHNEDECKRLRLMGVPEAP--SEAEKCALC 168
DB 231 ELQRAIKLKQTE-LDKLERLYRSPQNIPELFEILKILGIPASFSPIGVAEAFSAIS 309
OY 169 INDVFAVASEDMDSLTFGAPRFLRLIMPSSKIPVNEFDVAKYLEIELTMQFDLC 228
DB 310 ONNLAIVAAQDFDVLGSSMSINFLDNDNFHLPLQIMDPKRLAQLNLTFFGFDYC 369
OY 229 ILCCGCDVDSIKGIGOTALKLRHGSLESLLENIN-KDRQIPEPW--PYQEARLTK 285
DB 370 LMCGDFTSRIPKIGPYRALKLRIRYGNAPDVKALNVEKILITPTIKKFLTKKLT 429
OY 286 EPNVTLDIPELKWAPDEGLISFL 310
DB 430 D-----LPSNNELFSFI 441

RESULT 8
RAD2_YEAST STANDARD; PRT; 1031 AA.
AC P07276;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA repair protein RAD2.
GN RAD2 OR YGR258C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86083164; Pubmed=3000874;
RA Nicolet C.M., Chenevert J.M., Friedberg E.C.;
RT "The RAD2 gene of Saccharomyces cerevisiae: nucleotide sequence and
RT transcript mapping.";
RL Gene 36:225-234(1985).
RN (2)
RP REVISIONS. SEQUENCE FROM N.A.
RA MEDLINE=86223812; Pubmed=3011752;
RA Madura K., Prakash S.;

```

```

RT "Nucleotide sequence, transcript mapping, and regulation of the RAD2
RT gene of Saccharomyces cerevisiae.";
RL J. Bacteriol. 166:914-923(1986).
RN [3]
RP SEQUENCE FROM N.A.
RA Panzeri L., Agostoni Carbone M.L., Melchiorre P., Plevani P.,
RA Martegani E., Vancori M., Carignani G., Clemente M.L., Frontali L.,
RA Fabiani L., Marconi A., Ruzzi M., Sallola M.;
RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-51 FROM N.A.
RC STRAIN=5288C / FY1679;
RX MEDLINE=97245300; Pubmed=9090059;
RA Clemente M.L., Sartori G., Cardazzo B., Carignani G.;
RT "Analysis of an 11.6 kb region from the right arm of chromosome VII of
RT Saccharomyces cerevisiae between the RAD2 and the MSL1 genes reveals
RL the presence of three new genes.";
RN [5]
RP FUNCTION.
RX MEDLINE=94067324; Pubmed=8247134;
RA Habraken Y., Sung P., Prakash S.;
RT "Yeast excision repair gene RAD2 encodes a single-stranded DNA
RT endonuclease.";
RL Nature 366:365-368(1993).
CC -1- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN EXCISION
CC REPAIR OF DNA DAMAGED WITH UV LIGHT, BULKY ADDUCTS, OR CROSS-
CC LINKING AGENTS. ESSENTIAL FOR THE INCISION STEP OF EXCISION-
CC REPAIR.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M10275; AAB66928.1; -
DR EMBL; Y07777; CAA69080.1; -
DR EMBL; Z75043; CAA97287.1; -
DR PIR; A23960; A23960.
DR PIR; A29839; A29839.
DR SGD; S0003490; RAD2.
DR InterPro: IPR000513; Exo_N.1.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR001532; XPG_1.
DR Pfam; PF00752; XPG_1; 1.
DR Pfam; PF00867; XPG_N; 1.
DR PRINTS: PR00853; XGRADSUPER.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE; PS00841; XPG_1; 1.
DR PROSITE; PS00842; XPG_2; 1.
RW DNA repair; Nuclear protein; Hydrolase; Nuclease; Endonuclease;
KW Magnesium.
FT DOMAIN 1 95 N-DOMAIN.
FT DOMAIN 756 884 I-DOMAIN.
SQ SEQUENCE 1031 AA; 117837 MW; 682D4ACFBDF0F3 CRC64;

Query Match 12.0%; Score 232.5; DB 1; Length 1031;
Best Local Similarity 23.2%; Pred. No. 5.9e-08;
Matches 84; Conservative 67; Mismatches 18; Indels 53; Gaps 10;

OY 54 LTNAGEVYTHLQOMFNRTIRLLEAGIKV---YFDDKPPD-----MKQE---LAKR 101
DB 679 ISKEAEINDSRKVESINKER-----KPLIFDSEDEBDNIVEMIKKEQEBFDFTFKN 732

```

```

QY 102 YSKRDAATKDLTEAVEVGDK--DAIEKISKRTKVTROHNECKRLILMGVPEVAPSE 159
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 733 TLTLSAERNVAEENAFVDELFEEQOMKDKRSDSEVTMDMKIEVQELISRGIPYITAPME 792
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 160 AAEECAALCINDKPAFAVSEMDSLTFGAPFRLHILMDPSSKRPVMEFVAVLETEL 219
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 793 AEAQCAELLQMLVDGITTDDSDVFLFGGKTKYKNMF--HEKNVYEFYAESILKLLGL 849
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 220 TMDQFIDLCITCGCDYCDISIKGIGGOTALKLIRQHSIESILENLNDRYO----- 270
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 850 DAKNIEELAQILGSDYTNGLKMGVPVSSIEVIAEFGLNKNFKDWYNGGCDKRRQETENK 909
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 271 -----IPEDWPEYQEARLKFKEPNTLIDPELKTAPDEBELISFLVKDNG 315
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 910 FEKDLRKLVLNNEIILDDPSPVAVYDAYMRPEVDHDTTPFWGVPDLMLRSMFTOLG 969
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 316 F-----NEDRVTKALEIKISAKNKSOGRLSEFFKPIATYSAPLKRKETSPTSKAANK 371
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 970 WPEHKSDELLPLLRDVAKRKRRKKGKKRINEFF--PREYISOD-KKLNTSKRISTATGKXK 1027
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 372 TK 373
      :
DB 1028 KR 1029
      :
RESULT 9
XPG_HUMAN
ID XPG_HUMAN STANDARD; PRT; 1186 AA.
AC P28715;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA-repair protein complementing XP-G cells (Xeroderma pigmentosum
DE group G complementing protein) (DNA excision repair protein ERCC-5).
GN ERCC5 OR XPG OR XPGC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93247645; PubMed=8483504;
RA Scherly D., Nusspikel T., Corlet J., Ucla C., Bairoch A.,
RA Clarkson S.G.;
RT "Complementation of the DNA repair defect in Xeroderma pigmentosum
RT group G cells by a human cDNA related to yeast RAD2.";
RL Nature 363:187-185(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94173288; PubMed=7510366;
RA Shioml T., Hatada Y.-N., Saito T., Shioml N., Okuno Y., Yamazumi M.;
RA "An ERCC5 gene with homology to yeast RAD2 is involved in group G
RA xeroderma pigmentosum.";
RL Mutat. Res. 314:167-175(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94019313; PubMed=8413238;
RA MacInnes M.A., Dickson J.A., Hernandez R.R., Learmonth D., Lin G.Y.,
RA Mudgett J.S., Park M.S., Schaner S., Reynolds R.J., Strniste G.F.,
RA Yu J.Y.;
RT "Human ERCC5 cDNA-cosmid complementation for excision repair and
RT dipirite amino acid domains conserved with RAD proteins of
RT Saccharomyces cerevisiae and Schizosaccharomyces pombe.";
RL Mol. Cell. Biol. 13:6393-6402(1993).
RN [4]
RP SEQUENCE OF 1-88 FROM N.A.
RX MEDLINE=94373034; PubMed=8088606;
RA Samec S., Jones T.A., Corlet J., Scherly D., Sheer D., Wood R.D.,
RA Clarkson S.G.;
RT "The human gene for Xeroderma pigmentosum complementation group G
RT (XPG) maps to 13q33 by fluorescence in situ hybridization.";
RL Genomics 21:283-285(1994).

```

```

RP CHARACTERIZATION.
RX MEDLINE=94266772; PubMed=8206890;
RA O'Donovan A., Scherly D., Clarkson S.G., Wood R.D.;
RT "Isolation of active recombinant XPG protein, a human DNA repair
RT endonuclease.";
RL J. Biol. Chem. 269:15965-15968(1994).
RN [6]
RP CHARACTERIZATION.
RX MEDLINE=94376899; PubMed=8090225;
RA O'Donovan A., Davies A.A., McGee J.G., West S.C., Wood R.D.;
RT "XPG endonuclease makes the 3' incision in human DNA nucleotide
RT excision repair.";
RL Nature 371:432-435(1994).
RN [7]
RP CHARACTERIZATION.
RX MEDLINE=94359802; PubMed=8078765;
RA Habraken Y., Sung P., Prakash L., Prakash S.;
RT "Human Xeroderma pigmentosum group G gene encodes a DNA
RT endonuclease.";
RL Nucleic Acids Res. 22:3312-3316(1994).
RN [8]
RP REVIEW ON VARIANTS XP-G.
RX MEDLINE=99374920; PubMed=10447254;
RA Cleaver J.E., Thompson L.H., Richardson A.S., States J.C.;
RT "A summary of mutations in the UV-sensitive disorders: Xeroderma
RT pigmentosum, Cockayne syndrome, and trichothiodystrophy.";
RL Hum. Mutat. 14:9-22(1999).
RN [9]
RP VARIANT XP-G VAL-792, AND VARIANT ASP-1104.
RX MEDLINE=95038755; PubMed=7951246;
RA Nusspikel T., Clarkson S.G.;
RT "Mutations that disable the DNA repair gene XPG in a Xeroderma
RT pigmentosum group G patient.";
RL Hum. Mol. Genet. 3:963-967(1994).
RN [10]
RP VARIANT XP-G VAL-792.
RX MEDLINE=97250499; PubMed=9096355;
RA Nusspikel T., Lalle P., Leadon S.A., Cooper P.K., Clarkson S.G.;
RT "A common mutational pattern in Cockayne syndrome patients from
RT Xeroderma pigmentosum group G: implications for a second XPG
RT function.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:3116-3121(1997).
RN [11]
RP FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN DNA
RP EXCISION REPAIR. MAKES THE 3' INCISION IN DNA NUCLEOTIDE EXCISION
RP REPAIR.
RN [12]
RP SUBCELLULAR LOCATION: Nuclear (Probable).
RN [13]
RP DISEASE: COMPLEMENTS THE DEFECT IN XERODERMA PIGMENTOSUM (XP)
RP GROUP G (XP-G). AN AUTOSOMAL RECESSIVE DISEASE CHARACTERIZED BY
RP HYPERSENSITIVITY OF THE SKIN TO SUNLIGHT FOLLOWED BY HIGH
RP INCIDENCE OF SKIN CANCER AND FREQUENT NEUROLOGIC ABNORMALITIES.
RP DISEASE: ALSO INVOLVED IN COCKAYNE'S SYNDROME (CS); AN AUTOSOMAL
RP RECESSIVE DISEASE WHICH IS CHARACTERIZED BY A UV-SENSITIVE SKIN
RP (WITHOUT PIGMENTATION ABNORMALITIES), NEUROLOGICAL DYSFUNCTION
RP DUE TO DEMYELINATION OF NEURONS AND CALCIFICATION OF BASAL GANGLIA
RP (PSYCHOMOTOR RETARDATION, DEAFNESS, OPTIC ATROPHY, RETINAL
RP PIGMENTATION AND HYPERREFLEXES), AND DYSMORPHIC DWAFFISM (IMMATURE
RP SEXUAL DEVELOPMENT AND MICROCEPHALY).
RN [14]
RP SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
RP SUBFAMILY.
RN [15]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RP use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RP entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch).
RN [16]
RP EMBL, X69978; CAA49598.1; -
RP EMBL, D16305; BAA03812.1; -
RP EMBL, L20046; AAC37533.1; -
RP EMBL, X71341; CAA50481.1; -
RP EMBL, X71342; CAA50481.1; JOINED.

```

```

DR PIR: S35993; S35993.
DR MTM; 133530; -.
DR InterPro: IPR000513; Exo_N1.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR001532; XPG_1.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
DR PROSITE: PS00842; XPG_2; 1.
DR DNA repair: DNA-binding; Nuclear protein; Hydroxylase; Nuclease;
KW Endonuclease; Xeroderma pigmentosum; Cockayne's syndrome;
KW Polymorphism; Disease mutation.
FT DOMAIN 1 95
FT DOMAIN 753 881
FT DOMAIN 1057 1073
FT DOMAIN 254 254
FT VARIANT 792 792
FT VARIANT 1104 1104
FT VARIANT 1104 1104
FT CONFLICT 55 55
FT CONFLICT 120 122
FT CONFLICT 126 126
FT CONFLICT 264 266
FT CONFLICT 760 760
FT CONFLICT 796 796
FT CONFLICT 864 872
FT CONFLICT 959 959
FT CONFLICT 1186 AA; 133328 MW; 8702EC68E080F1C4 CRC64;
SQ SEQUENCE

Query Match 11.4%; Score 220.5; DB 1; Length 1186;
Best Local Similarity 21.3%; Pred. No. 4.3e-07;
Matches 94; Conservative 71; Mismatches 149; Indels 127; Gaps 15;
15;
15 PRAMEQKESYFGRIADVADSMISYQFLIVGRTGMETLNEAGEVTSHLQGMFNRTIR 74
635 PRAVEMELDS-----ESESQSGSFLEV-----QSVISDE-----ELQAEPPET-- 673
75 LLEAGIKPVYVVDGKPPMKKOE-----LAKRSKRDATRKDLLEAVEV 118
674 -----SKPPSGEGEEELVGTREGAPAESESLLRDNSEKDVDDGEPOEA--- 717
119 GQKDAIEKL-----SKRTVYTRQHN-----DCKRLL 146
718 -EKDADDSLHEMODINLELEETLESNLAQONSIAKQKQGERIAVATYGOMFLESQELL 776
147 RLMGVAVNEVESEAEACALCINDKVFVASEMDSLTFGAPRFLRLHMDSSKRTIPM 206
777 RLFGIPIYIADPMEAEACALILDLTDQTSGITDSDIMLFGARHYRNFE--NNKTFVE 833
207 EEPVNAVLELELTMDQFIDLITLGCDCDSIKIGISGOTAKLIRQ--HGSISILE- 262
834 YIYGVDFHNDGLDRKRLINLAIVLSDTTEGIPYGCYTAHEIINFEFGHG-LEPLKLF 892
263 -----NLNKD--RYQIPEDMPYQEARLFEKPNVTLDIPELKWTA 300
893 SEWMHEAQKPKIRPNPHDKVKKKLRITLQIFGFPNPVAAVEYLRKPVDDSKSGSLWKG 952
301 PDEGLISFLVKNQNGENEDRYTKAI-EKIKSAKNSKSGRLEFEFPRTATTSAPLRKRET 359
953 PDIDKIREFCORFYGMWRKRTKTDSELPVLRKQDADQOTLRIDSEFRLAQOEKEDAKRIKS 1012
360 S-----DKSKAANK 370
1013 QRLNRAVTCMLRKREKKAASE 1033

```

```

RESULT 10
XPG_MOUSE STANDARD; PRT: 1170 AA.
AC P35689; 061528; 064248;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA-repair protein complementing XP-C cells homolog (Xeroderma
DE pigmentosum group G complementing protein homolog) (DNA excision
DE repair protein ERCC-5).
GN ERCC5 OR XPG OR ERCC-5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9417328; PubMed=7510366;
RA Shiomoto T., Harada Y.-N., Saito T., Shiomoto N., Okuno Y., Yamazumi M.;
RT "An ERCC5 gene with homology to yeast RAD2 is involved in group G
RT xeroderma pigmentosum."
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/10; TISSUE=Liver;
RA Harada Y.N., Matsuda Y., Shiomoto N., Shiomoto T.;
RT "Complementary DNA sequence and chromosomal localization of xpg, the
RT mouse counterpart of human repair gene XPG/ERCC5."
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=DNA/2;
RA Ludwig D.L., Mudgett J.S., Park M.S., Perez-Castro A.V.,
RT "Molecular cloning and structural analysis of the functional mouse
RT genomic XPG gene."
RN Mamm. Genome 7:644-649(1996).
CC -1- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN DNA
CC EXCISION REPAIR. MAKES THE 3' INCISION IN DNA NUCLEOTIDE EXCISION
CC REPAIR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D16306; BAA03813.1; -
DR EMBL; U40796; AAB17885.1; -
DR EMBL; U40795; AAB17885.1; JOINED.
DR EMBL; U39892; AAB17885.1; JOINED.
DR EMBL; U39893; AAB17885.1; JOINED.
DR EMBL; U39894; AAB17885.1; JOINED.
DR EMBL; U39895; AAB17885.1; JOINED.
DR EMBL; U40073; AAB17885.1; JOINED.
DR EMBL; U40431; AAB17885.1; JOINED.
DR EMBL; U40432; AAB17885.1; JOINED.
DR EMBL; U40668; AAB17885.1; JOINED.
DR EMBL; U40669; AAB17885.1; JOINED.
DR EMBL; U40670; AAB17885.1; JOINED.
DR EMBL; U40792; AAB17885.1; JOINED.
DR EMBL; U40793; AAB17885.1; JOINED.
DR EMBL; U40794; AAB17885.1; JOINED.
DR MGD; MGI:103582; ERCC5.
DR InterPro: IPR000513; Exo_N1.
DR InterPro: IPR003584; HNH_2.

```

```

DR InterPro: IPR001532; XPG_I.
DR Pfam: PF00867; XPG_I.1.
DR Pfam: PF00752; XPG_N.1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HNH2.1.
DR SMART: SM00484; XPGI.1.
DR SMART: SM00485; XPGN.1.
DR PROSITE: PS00841; XPG_1.1.
DR PROSITE: PS00842; XPG_2.1.
DR DNA repair: DNA-binding; Nuclear protein; Hydrolase; Nuclease;
KM Endonuclease.
FT DOMAIN 1 95 N-DOMAIN.
FT DOMAIN 752 882 I-DOMAIN.
FT DOMAIN 1049 1065 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT VARIANT 388 388 R->C (IN STRAIN DBA/2).
FT VARIANT 488 488 R->C (IN STRAIN DBA/2).
FT VARIANT 688 688 I->T (IN STRAIN DBA/2).
FT VARIANT 1015 1015 S->N (IN STRAIN DBA/2).
FT VARIANT 1021 1021 S->I (IN STRAIN DBA/2).
FT VARIANT 1121 1121 S->P (IN STRAIN DBA/2).
FT CONFLICT 227 227 N->M (IN REF. 1).
FT CONFLICT 249 249 N->M (IN REF. 1).
FT CONFLICT 300 302 VMD->MDE (IN REF. 1).
FT CONFLICT 313 313 N->M (IN REF. 1).
FT CONFLICT 320 320 N->M (IN REF. 1).
FT CONFLICT 399 399 N->M (IN REF. 1).
FT CONFLICT 408 408 D->DVQGVG (IN REF. 1).
FT CONFLICT 581 581 N->NSASEVIGPV (IN REF. 1).
FT CONFLICT 795 795 M->V (IN REF. 1).
FT CONFLICT 1039 1039 A->AMERFEL (IN REF. 1).
FT CONFLICT 1134 1134 S->SD (IN REF. 1).
FT CONFLICT 1157 1158 KL->RR (IN REF. 1).
SQ SEQUENCE 1170 AA; 130864 MW; 4058F07FFD1770ED CRC64;

Query Match 11.3%; Score 220; DB 1; Length 1170;
Best Local Similarity 21.6%; Pred No. 4.6e-07;
Matches 88; Conservative 71; Mismatches 156; Indels 92; Gaps 13;

QY 13 NAKRANKKQKFESEYFGRKIAVDASMSYFLIVGRTGKMTLTNAGEVTSHL-----65
DB 633 SAKRPMQPMEMES-----ESESSEDSGSEFIEVOSVNSSELOTSSASTHLSKDAEE 684
QY 66 -----OGMNFRTLLLEAGIKRVYVFQDKPPDMKKKQKELAKRYSRDATKDLT---EAY 116
DB 685 PRELEGETSRQTECLIQ-----DSSDIEMEGHRADDDAEDMPMEMODI 730
QY 117 EVGDKDAIEK--LSKRYKVTROHNE-----DCKRLRLMGVPEVPEAE 161
DB 731 NLEELDALESNLLAEONSILKAKQODRIAASVTGOMFLESQELLRLFGVPIQAPMEAE 790
QY 162 AECAALCINDKFAVAASEMDSLTFGARFLRLHMDPSKRIPWNE-FDYAKVLELELT 220
DB 791 AOCAMLDLSQTSQTTTIDSDIWLFGA---HHVYKNFNNKKFVEYQYVDYFSQGLID 846
QY 221 MDGFIDLCILGGDYCDYSIKIGOGATLKLIRQH--GSIESIL-----ENLNK---266
DB 847 RNKLINLAYLLGSDYTEGIEPIYVGCYVAETILNEFGRGDLPLKSEWHEKONKKYAE 906
QY 267 -----DRYOIPEDMPYOEARLKEFKNVTLIDIPELKTPADEGLISFLYKNG 315
DB 907 NPVDTKWKKKLRLQLLPGFNPNAVADALRPVDDSRGSLMKGPDVDKREFQYRFG 966
QY 316 FNEDRVTKATEKI-KSAKNSSGRLESEFFKPTATTSAPLKRKETS 361
DB 967 WNRKKTDESLEYVILKHLNAHOTLRIDSFRR-----LAQGEROD 1005

RESULT 11
EXOT_YEAST STANDARD: PRT; 702 AA.
AC P39875;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)

```

```

DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Exonuclease I (EXO I) (DHSI protein).
EX EXOI OR DHSI OR YOR03C OR OR26.23.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycs.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 220-702 FROM N.A.
RX MEDLINE=94169519; PubMed=7764548;
RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
RT hyper-sensitive mutation of a gene, DHSI, which complements a drug-;
RL bioosci. Biotechnol. Biochem. 58:391-395(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RA de Haan M., Maarse A.C., Griwell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Landt O., Hiesel R., Unseld M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Tishkoff D.X., Boerger A.L., Filosi N., Gaida G.M., Bertrand P.,
RA Kane M.F., Kolodner R.D.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP FUNCTION.
RX MEDLINE=97265410; PubMed=9111347;
RA Florentini P., Huang K.N., Tishkoff D.X., Kolodner R.D.,
RA Symington L.S.;
RT "Exonuclease I of Saccharomyces cerevisiae functions in mitotic
RT recombination in vivo and in vitro.";
RL Mol. Cell. Biol. 17:2764-2773(1997).
CC -I- FUNCTION: 5'->3' double-stranded DNA exonuclease involved in
CC mismatch repair and eventually also in mitotic recombination
CC between direct repeats.
CC -I- COFACTOR: Magnesium or manganese.
CC -I- ENZYME REGULATION: Inactivated by calcium and zinc ions.
CC -I- SUBUNIT: INTERACTS WITH MISMATCH REPAIR PROTEIN MSH2.
CC -I- SUBCELLULAR LOCATION: Nuclear (potential).
CC -I- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S69345; AAC60570.1; -
DR EMBL; X87331; CAAG0749.1; -
DR EMBL; Z74941; CAAG9223.1; -
DR EMBL; U084134; AAB47428.1; -
DR PIR; JC2068; JC2068.
DR HSSP; Q29075; TNKI.
DR SGD; S0005559; DHSI.
DR InterPro: IPR000513; EXO_N_I.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR001532; XPG_I.
DR Pfam: PF00867; XPG_I.1.
DR Pfam: PF00752; XPG_N.1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HNH2.1.
DR SMART: SM00484; XPGI.1.
DR SMART: SM00485; XPGN.1.
DR PROSITE: PS00841; XPG_1.1.
DR PROSITE: PS00842; XPG_2.1.1.
DR DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.
FT DOMAIN 1 96 N-DOMAIN.

```

FT DOMAIN 114 247 I-DOMAIN.  
 FT DOMAIN 501 520 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 537 553 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 618 625 ASP-RICH (ACIDIC).  
 SO SEQUENCE 702 AA; 80162 MW; 7822F6B265DB3AA CRC64;

Query Match 11.2%; Score 218; DB 1; Length 702;  
 Best Local Similarity 23.7%; Pred. No. 3.3e-07;  
 Matches 99; Conservative 70; Mismatches 176; Indels 72; Gaps 14;

OY 1 MGIG-LTKLLADNAPKMKQKFESEFGKRIADASMSIYO-----FLIVGRTGEMT 53  
 DB 1 MGIGLTPQLKPIQNPVSLR-----YEGVLADGYAMHRAACCAVEIAMGK----- 50  
 OY 54 LTNEGEVTSHLQGMFNFTIRLEAGIKPVYFEDGKPPDMKQKOLARYSKRDATKDLT 113  
 DB 51 -----PDKTLQPIKRFSLTKTKVPRYPVFDGDAIPVKKSKRROKRKE-NKAIA 103  
 OY 114 EAV-EVGDKAIEKLSKRTVKTROHNEDECKRLRLMGVPEVAPSEAEACALCTNDK 172  
 DB 104 ERLMAGCKKNAMDYFQKCVDPITPEMAKCIICCKLNGIRYIVAFEDSQMVELEQKNI 163  
 OY 173 VFAVASEMDSLTFGAPRFLRLHMDPSS-----KTIPEVMEFVAKVLELEL 219  
 DB 164 VOGIISESDSLVFGCRRLTKLNDYGCLECRDNFTKLEKKEPLG-----SL 212  
 OY 220 TMDQFIDLCICGDCYCSIKIGGOTALKLRHGSIESILENLNKD-RVQIPEDW--P 276  
 DB 213 TNEEITWYCSGCDYTNIGIPKVLITAMKLVRRNTERITLSIORGKLMDPTKYNE 272  
 OY 277 YOEARRLEKRPVY-----TLDIPELKWTADEGLSLFVKNQGFEDVNTVAIEKIK 329  
 DB 273 YEAAVLAFQFQFVCPKIRKRIYSLNEIPLYLNDTSKRRLYLACIGFYIHETQKKQIVH 332  
 OY 330 SAKN-----KSSQGRLE--SFFKPTATTSAPLKRKETS-----KTSKAANKTK 373  
 DB 333 FDDDDHHLHLTAQGDLPNYPFHQPLANREHKLQLASNSNIEFGKTTNTNSEAKVK 389

RESULT 12  
 IDIN7 YEAST STANDARD; PRT; 430 AA.  
 AC 012086;  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE DNA-damage inducible protein DIN7.  
 GN DIN7 OR DIN3 OR YDR263C OR YD9320B.02C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C / AB972;  
 RA Murphy L., Harris D., Bartrell B.G., Rajandream M.A., Walsh S.V.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: 5'->3' double-stranded DNA exonuclease (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -1- INDUCTION: By UV light, methyl methane-sulfonate (MMS) or hydroxyurea (HU), and during meiosis.  
 CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: X90707; CA62233.1; -  
 DR EMBL: 270202; CA94102.1; -  
 DR EMBL: 268290; CA92581.1; -  
 DR SGD: S0002671; DIN7.  
 DR InterPro: IPR005013; EXO\_N\_1.  
 DR InterPro: IPR003584; HH2.  
 DR InterPro: IPR001532; XPG\_1.  
 DR Pfam: PF00867; XPG\_1; 1.  
 DR Pfam: PF00752; XPG\_1; 1.  
 DR PRINTS: PR00853; XPGRADSUPER.  
 DR SMART: SM00279; HH2; 1.  
 DR SMART: SM00484; XPG1; 1.  
 DR SMART: SM00485; XPGN; 1.  
 DR PROSITE: PS00841; XPG\_1; 1.  
 DR PROSITE: PS00842; XPG\_2; 1.  
 KW DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.  
 FT DOMAIN 114 247 N-DOMAIN.  
 FT DOMAIN 1 96 I-DOMAIN.  
 SO SEQUENCE 430 AA; 49034 MW; 2BC23D30832C79E9 CRC64;

Query Match 10.5%; Score 204.5; DB 1; Length 430;  
 Best Local Similarity 24.7%; Pred. No. 1.4e-06;  
 Matches 68; Conservative 53; Mismatches 129; Indels 25; Gaps 6;

OY 1 MGIGLTKLLADNAPKMKQKFESEFGKRIADASMSIYO-----FLIVGRTGEMT 54  
 DB 1 MGIGLTPQLKPIQNPVSLR-----YEGVLADGYAMHRAACCAVEIAMGK----- 51  
 OY 55 TNEAGEVTSHLQGMFNFTIRLEAGIKPVYFEDGKPPDMKQKOLARYSKRDATKDLT 114  
 DB 52 TNK-----YLOFEIKRLQILKRLKIPYVFDGSLFVANHETRRKKRLLENIAK 105  
 OY 115 AVEVGDKDAIEKLSKRTVKTROHNEDECKRLRLMGVPEVAPSEAEACALCTNDKYE 174  
 DB 106 LMSAGNRYNAMEYFQKSVDPITPEMAKCIIDYCKLHSIPYIVAFEDPQVYLEKMGILQ 165  
 OY 175 AVASEMDSLTFGAPRFLRLHMDPSSKRIYMEPDVAKVLEL-----ELTMDQFIDLCIC 231  
 DB 166 GIISEDLSLVFGCKTLITKRLND-QGKALEISKDFSALEPNPGLSLSPQFRNLVCLA 224  
 OY 232 GCDYCSIKIGGOTALKLRHGSIESILENLNK 266  
 DB 225 GCDYTGIMKVGVYTKIKYRSEMDLILQIER 259

RESULT 13  
 ID RA13\_SCHPO STANDARD; PRT; 1112 AA.  
 AC P28706; O59728;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE DNA repair protein rad13.  
 GN RAD13 OR SPB387.08C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=93219111; PubMed=8464724;  
 RA Carr A.M., Sheldrick K.S., Murray J.M., Al-Harithy R., Watts F.Z.,  
 RA Lehmann A.R.;  
 RL "Evolutionary conservation of excision repair in Schizosaccharomyces  
 RT pombe: evidence for a family of sequences related to the

```

RT Saccharomyces cerevisiae RAD2 gene.;
RL Nucleic Acids Res. 21:1345-1349(1993).
RN
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Lyne M., Wood V., Rajandream M.A., Barrell B.G., Brown D.,
RA Churcher C.M.;
RL Submitted (MAY-1998) to the EMBL/GenBank/CDDB databases.
CC -!- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN EXCISION
CC REPAIR OF DNA DAMAGED WITH UV LIGHT, BULKY ADDUCTS, OR CROSS-
CC LINKING AGENTS. ESSENTIAL FOR THE INCISION STEP OF EXCISION-
CC REPAIR (PROBABLE).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X66795; CAA47291.1; -.
DR EMBL: AL023534; CAA19011.1; -.
DR PIR: S22862; S22862.
DR PIR: S30301; S30301.
DR InterPro: IPR000513; Exo_N_I.
DR InterPro: IPR001191; Gemini_AL1.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR003903; UIM.
DR InterPro: IPR001532; XPG_I.
DR Pfam: PF00799; Gemini_AL1.1.
DR Pfam: PF00867; XPG_I.1.
DR Pfam: PF00752; XPG_N.1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00484; XPGI; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
DR PROSITE: PS00842; XPG_2; 1.
DR DNA repair; Nuclear protein; Hydrolyase; Nuclease; Endonuclease.
FT DOMAIN 1 870 N-DOMAIN.
FT CONFLICT 742 870 D->N (IN REF. 1).
FT CONFLICT 738 743 LKMKR->AQKSKRG (IN REF. 1).
SQ SEQUENCE 1112 AA; 126328 MW; 7ECF4229D5BF4768 CRC64;

Query Match 10.5%; Score 204.5; DB 1; Length 1112;
Best Local Similarity 21.4%; Pred. No. 4.5e-06;
Matches 88; Conservative 68; Mismatches 148; Indels 107; Gaps 15;

```

```

OY 254 -----HGSTES---ILENLK--DRQIPEDMPYQEARLKEPNVTLDIPELKW 298
DB 880 KMFQRLSTGHASKNDVNTVPVKRRKILVGLTLPSEFPNPLVDAYLHPVADDSKOSFOW 939
OY 299 TAPDEGLISFLVKNQGFEDNRTKAIEKISAKK-----SSQGRLSFPR 345
DB 940 GTPDELEHGFMLATYGMWQSKQRTNEVLPLVTDHMKKQFVGTQGNLQFPE 990

```

## RESULT 14

XPG\_XENLA STANDARD; PRT; 1196 AA.

```

ID XPG_XENLA
AC P14629;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA-repair protein complementing XP-G cells homolog (Xeroderma
DE pigmentosum group G complementing protein homolog).
GN ERKCS OR XPG OR XPGC.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93247645; PubMed=8483504;
RA Scherly D., Nusspikel T., Corlet J., UCLA C., Bairoch A.,
RA Clarkson S.G.;
RT "Complementation of the DNA repair defect in Xeroderma pigmentosum
RT group G cells by a human cDNA related to yeast RAD2."
RL Nature 363:182-185(1993).
CC -!- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN DNA
CC EXCISION REPAIR. MAKES THE 3' INCISION IN DNA NUCLEOTIDE EXCISION
CC REPAIR (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69977; CAA49597.1; -.
DR PIR: S35994; S35994.
DR InterPro: IPR000513; Exo_N_I.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR001532; XPG_I.
DR Pfam: PF00867; XPG_I.1.
DR Pfam: PF00752; XPG_N.1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00484; XPGI; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
DR PROSITE: PS00842; XPG_2; 1.
DR DNA repair; DNA-binding; Nuclear protein; Hydrolyase; Nuclease;
FT DOMAIN 1 95 N-DOMAIN.
FT DOMAIN 786 914 I-DOMAIN.
FT CONFLICT 936 952 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT CONFLICT 1079 1095 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 1196 AA; 134206 MW; 1F1CE1891A3C0623 CRC64;

```

```

Query Match 10.3%; Score 200; DB 1; Length 1196;
Best Local Similarity 21.9%; Pred. No. 9.9e-06;
Matches 72; Conservative 64; Mismatches 137; Indels 56; Gaps 11;

```

```

OY 88 GKPEPMKKQELAKRYSKRDATKIDLTAEVVGDKDAIEK---LSKRYVYKTRQNE--- 140

```

```

Db 740 GEPHEOSNSESK---DLDDVSNEMOD-1SYDELESLNNLYVOOTS1QAOOQOQOEETIA 795
OY 141 -----DEKRLRLMGVPEVPEAPSEAEACALCINDKFAVAASEDMSTFTGAPRL 192
Db 796 TVYQOMKLEQDGLQGLFGIPYAPPEAEACQALIDLDQISGTTTDDSDIWMFGAHVY 855
OY 193 RHLMDSSKATPYMEFDVAKVLEELTMDOFIDLCICGCDYCDISKIGIGGQALKLIR 252
Db 856 KNFP---SOKRHEVYOVADINHQGLDRSKLILNAYLIGSDYIEGIPYGVYSAMEILN 912
OY 253 OH--GSIESILE-----NINMDRYOITP-----EDMPYOEARLKEP 287
Db 913 EFPQGLPELVKFKEMWSEAKOKKMRPNPNDTKVKKRLRLIDQSPNPVAVASAYLKP 972
OY 288 NVTLIDPELWTPAPDEGLISFLVKDNGF---NEDRYKALEIKTSKANKSSQGRLESFF 344
Db 973 VVDESKAFSFGRPDLQJREFCESFEGYRLKTDVLLPV--IKQLNAQOTQLRIDSF 1030
OY 345 KPATTSAPLKRKRETSDDKTSKAAANKTK 373
Db 1031 RLEQHEAAGLK-----SQRLEBAVTCMKRK 1055

RESULT 15
DPO1_THECA STANDARD: PRT: 834 AA.
AC P80194;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA polymerase I, thermostable (EC 2.7.7.7) (TMC polymerase 1).
GN POLA.
OS Thermus aquaticus (subsp. caldophilus).
OC Bacteria; Thermus/Delnococcus group; Thermus group; Thermus.
OX NCBI_TaxID=272;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GR24;
RX MEDLINE=93285135; PubMed=8508785;
RA Park J.H., Kim J.S., Kwon S.-T., Lee D.-S.;
RT "Purification and characterization of Thermus caldophilus GK24 DNA
   polymerase".
RL Eur. J. Biochem. 214:135-140(1993).
CC -I- FUNCTION: THERMOSTABLE ENZYME THAT HAS 5'-TO-3' EXONUCLEASE
CC ACTIVITY AND NO 3'-TO-5' EXONUCLEASE ACTIVITY.
CC -I- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -I- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U62584; AAB81398.1;
CC PIR: S33287; S33287.
CC HSSP: P19821; IMAO.
DR InterPro: IPR002421; 5_3_exonuclease.
DR InterPro: IPR002296; DNA_polI.
DR InterPro: IPR001098; DNA_pol_A.
DR InterPro: IPR000513; Exo_N_I.
DR InterPro: IPR003583; HHH_1.
DR InterPro: IPR003584; HHH_2.
DR Pfam: PF01367; 5_3_exonuclease; 1.

```

```

DR Pfam: PF02739; 5_3_exonuc_N_1.
DR Pfam: PF00476; DNA_pol_A_1.
DR PRINTS: PR00868; DNAPOLI.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00278; HHH1; 1.
DR SMART: SM00279; HHH2; 1.
DR SMART: SM00482; POLAC; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
KM Transferrase, DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding.
FT DOMAIN 412 834 POLYMERASE (BY SIMILARITY).
SQ SEQUENCE 834 AA; 93798 MW; A851FF3C3076348E CnC64;

Query Match 9.5%; Score 184.5; DB 1; Length 834;
Best Local Similarity 25.4%; Pred. No. 6; 6e-05;
Matches 78; Conservative 47; Mismatches 115; Indels 67; Gaps 13;

OY 28 GRKIAVDASMSITQFLIVGRTGMEITLNEAGEVTSHQGMNRTIRLL-EGATKPYV-Y 85
Db 12 GRVLVGGHHLAAYTFE-----ALKGLTTSKGEVQAVYGEFAKSLKALKEDGYKAVFV 66
OY 86 FDGKPPMKKQELAKRYSKRDKDITKDLTEAVVGVCDKDALEKLSKRTVAVTQHNEDCKRL 145
Db 67 FDAKAPGFRHEAV-----FAKAKGRAPTPEDPPROLALI-----KEL 103
OY 146 IRLMGVPEVPEAPSEAEACALCINDKV-----FAVASEDMDSLTGCAPRFLR--HL 195
Db 104 VDLGFTFRLEVPGEADVDVATLATLAKNPEKEGYEVRLITADRLDQVSDRAVVLHPGHL 163
OY 196 MDPSKSKATPYMEFDVAKVLEELTMDOFIDLCICG--CDYCDISKIGIGGQALKLIRQ 253
Db 164 ITPB-----WLMOKYGLKPEQWVDFRALGVDESDLPVKGKIGKTKALKLKE 211
OY 254 HGSTESILENINKRVOIPED-----WPYOEARLKEP-----PVTLDIDPELWTPAPDE 303
Db 212 WGSLEMLINKNDRVK---PENVEKIKAHLEDRLSLSLSRVTRTDPLFVLDLAGREPR 268
OY 304 EGLISFL 310
Db 269 EGLRAFL 275

Search completed: November 5, 2002, 16:28:18
Job time : 19 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:26:10 ; Search time 26 seconds  
(without alignments)  
2521.734 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939  
Sequence: 1 MGIGLTKRLADNAPKAKE.....SDKTSKAAANKTKAGKKK 379

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_prodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1807.5	93.2	380	10	09SX06
2	1313.5	67.7	362	10	065251
3	1069	55.1	382	13	P70040
4	1068	55.1	382	13	057351
5	1068	55.1	382	13	P70054
6	1063.5	54.8	381	13	090YB0
7	1043	53.8	380	11	091Z50
8	1040	53.6	380	11	09JH7
9	1036.5	53.5	382	5	09N372
10	1009	52.0	378	11	099M86
11	998.5	51.5	385	5	096830
12	887	45.7	650	5	09G201
13	887	45.7	672	5	09UOK1
14	648.5	33.4	340	1	093634
15	636	32.8	343	17	09V0B9
16	632.5	32.6	343	17	050123

17	542	28.0	328	17	027670	027670 methanother
18	538.5	27.8	336	17	029975	029975 archaeoglob
19	519	26.8	304	17	0976H6	0976H6 sulfolobus
20	514.5	26.5	401	17	09YFY5	09YFY5 aeropyrum p
21	506	26.1	302	17	096B08	096B08 sulfolobus
22	468	24.1	335	17	09H4D4	09H4D4 thermoplasm
23	446	23.0	335	17	097B98	097B98 thermoplasm
24	369	19.0	327	17	09H027	09H027 halobacteri
25	336	17.3	316	5	096154	096154 plasmodium
26	278	14.3	415	12	091FF5	091FF5 chilo iride
27	267.5	13.8	317	10	09G7N8	09G7N8 arabidopsis
28	257	13.3	726	5	09VR00	09VR00 drosophila
29	257	13.3	726	5	09U906	09U906 drosophila
30	243.5	12.6	1257	5	0917N6	0917N6 drosophila
31	241	12.4	298	13	09QTF3	09QTF3 red sea bre
32	240	12.4	734	13	09W6K2	09W6K2 xenopus lae
33	235	12.1	497	10	09LEP2	09LEP2 arabidopsis
34	230	11.9	837	11	09Q021	09Q021 mus musculu
35	230	11.9	837	11	0923A5	0923A5 mus musculu
36	227	11.7	800	4	075466	075466 homo sapien
37	227	11.7	803	4	060545	060545 homo sapien
38	227	11.7	803	4	075214	075214 homo sapien
39	227	11.7	846	4	09UNW0	09UNW0 homo sapien
40	227	11.7	846	4	09UQ84	09UQ84 homo sapien
41	227	11.7	846	4	0961J1	0961J1 homo sapien
42	220.5	11.4	1186	4	09HD59	09HD59 homo sapien
43	219.5	11.3	1479	10	09ART5	09ART5 arabidopsis
44	219.5	11.3	1522	10	09LRT2	09LRT2 arabidopsis
45	211	10.9	732	5	024557	024557 drosophila

#### ALIGNMENTS

RESULT 1	09SX06	PRELIMINARY:	PRT:	380 AA.
ID	09SX06			
AC	09SX06			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	FEN-1.			
GN	OSFEN-1.			
OS	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=4530;			
RN	[1]			
RA	SEQUENCE FROM N.A.			
RA	MEDLINE=20256470; PubMed=10798612;			
RA	Kimura S., Ueda T., Hatanaka M., Takenouchi M., Hashimoto J.,			
RA	Sakaguchi K.;			
RT	"Plant homologue of flap endonuclease-1: molecular cloning,			
RT	characterization, and evidence of expression in meristematic			
RT	tissues.";			
RL	Plant MOL. Biol. 42:415-427(2000).			
DR	EMBL: AB021666; BAA36171.1; .			
DR	HSSP: 058839; 1A76.			
DR	InterPro: IPR002421; 5_3-exonuclease.			
DR	InterPro: IPR003584; HHH_2.			
DR	InterPro: IPR001532; XPG_1.			
DR	Pfam: PF01367; 5_3-exonuclease; 1.			
DR	Pfam: PF00867; XPG_I; 1.			
DR	Pfam: PF00752; XPG_N; 1.			
DR	PRINTS: PRO0853; XPGRADSUPR.			
DR	SMART: SM00475; 53EXoc; 1.			
DR	SMART: SM00279; Hhh2; 1.			
DR	SMART: SM00484; XPGI; 1.			
DR	SMART: SM00485; XPGN; 1.			
DR	PROSITE: PS00841; XPG_1; 1.			
SO	SEQUENCE 380 AA; 42792 MW; E0148AFA95A7364 CRC64;			

Query Match 93.2%; Score 1807.5; DB 10; Length 380;  
 Best Local Similarity 92.1%; Pred. No. 7, 2e-121;  
 Matches 350; Conservative 16; Mismatches 13; Indels 1; Gaps 1;

QY 1 MGIGLTKLLADNAPKAMKEOKESFYGRKRIADVASKSYOFLIVGRTGMETLINEAGE 60  
 D 1 MGIGLTKLLADNAPKAMKEOKESFYGRRIADVASKSYOFLIVGRTGMETLINEAGE 60  
 QY 61 VTSHQGFNFTIRLEAGIKPVYFPGKPPDMKQOELAKRYSKRQDADKDLTEAVEGD 120  
 D 61 VTSHQGFNFTIRLEAGIKPVYFPGKPPDLKKOELAKRYSKRQDADKDLTEAVEGD 120  
 QY 121 KDAIEKSKRTVKTROHNEDECKRLRLMGVPVEAPSEAEACALCINDKVEAVASD 180  
 D 121 KDAIEKSKRTVKTROHNEDECKRLRLMGVPVEAPSEAEACALCINDKVEAVASD 180  
 QY 181 MDLTFGAPRFLRLHMDPPSSKKTIPVMEFVAKVLELELMDQFIDLCILGCDYCDSTK 240  
 D 181 MDLTFGAPRFLRLHMDPPSSKKTIPVMEFVAKVLELELMDQFIDLCILGCDYCDSTK 240  
 QY 241 GIGGOTALKLIRHGSIESILENINMDRYOIPEDWPYQEARLKEPNVTLDIPELKWTA 300  
 D 241 GIGGOTALKLIRHGSIESILENINMDRYOIPEDWPYQEARLKEPNVTLDIPELKWTA 300  
 QY 301 PDEGLISFLVKNFGNPDRTYKAIKIKSAKSSSGRLSEFPEKPTATSPARKETS 360  
 D 301 PDEGLISFLVKNFGNPDRTYKAIKIKSAKSSSGRLSEFPEKPTATSPARKETS 360  
 QY 361 DKTSKAANKKTK-AGSKKK 379  
 D 361 EKPTKAVANKKTKGAGCKK 380

RESULT 2  
 ID 065251 PRELIMINARY; PRT; 362 AA.  
 AC 065251;  
 DT 01-AUG-1998 (Tremblrel. 07, Created)  
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE F2IE10.3 PROTEIN.  
 GN F2IE10.3.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Davidson S., Kohlring T., David M., O'Brian D.;  
 RT "The sequence of A. thaliana F2IE10."  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Washu.  
 RT Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 RN 13  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Waterston R.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 RN 14  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Wilson R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF058914; AAC13596.1; -  
 DR HSSP: Q58839; 1A76.  
 DR InterPro: IPR002421; 5\_3\_exonuclease.

DR InterPro: IPR000513; Exo\_N\_I.  
 DR InterPro: IPR003584; HHN\_2.  
 DR InterPro: IPR001532; XPG\_I.  
 DR Pfam: PF01367; 5\_3\_exonuclease; 1.  
 DR Pfam: PF00867; XPG\_1; 1.  
 DR Pfam: PF00752; XPG\_N; 1.  
 DR PRINTS: PR00853; XPGADSUPER.  
 DR SMART: SM00279; HHN2; 1.  
 DR SMART: SM00484; XPG1; 1.  
 DR SMART: SM00485; XPGN; 1.  
 DR PROSITE: PS00841; XPG\_1; 1.  
 SO SEQUENCE 362 AA; 41205 MW; A4DD1706C3AD29D CXC64;

Query Match 67.7%; Score 1313.5; DB 10; Length 362;  
 Best Local Similarity 72.1%; Pred. No. 1e-85;  
 Matches 259; Conservative 30; Mismatches 33; Indels 37; Gaps 3;

QY 1 MGIGLTKLLADNAPKAMKEOKESFYGRKRIADVASKSYOFLIVGRTGMETLINEAGE 60  
 D 1 MGIGLTKLLADNAPSCMKOEKFESEFGRKIAVDASKSYOFLIVNAPSSPL-LLEELGL 59  
 QY 61 VTS-----HLOGFNFTIRLEAGIKPVYVF 86  
 D 60 KCSLMKLVKSLAWIFLCDVLCIVFVYVYIWMFNHLOGFNFTIRLEAGIKPVYVF 119  
 QY 87 DGRPPDMKQOELAKRYSKRQDADKDLTEAVEGDKALIEKSKRTVKTROHNEDECKRL 146  
 D 120 DGRPELKEKQELAKRYSKRQDADKDLTEAVEGDIEGKNEDEKYSKRTVKTROHNEDECKRL 179  
 QY 147 RLMGVPVEAPSEAEACALCINDKVPFAVASEDMDSLTFGAPRFLRLHMDPPSSKKTIPVM 206  
 D 180 RLMGVPVEAPSEAEACALCCKSGKVYGVASEDMDSLTFGAPRFLRLHMDPPSSKKTIPVM 239  
 QY 207 EFDYAKVLELELTMQFIDLCILGCDYCDSTIKGIGGOTALKLIRHGSIESILENIN 266  
 D 240 EFEVAKLIEELOLTMQFIDLCILGCDYCDSTIRGIGGOTALKLIRHGSIESILENIN 299  
 QY 267 DRYOIPEDWPYQEARLKEPNVTLDIPEL--KMTAPDEGLISFLVKNFGNPDRTYK 323  
 D 300 ERYOIPEDWPYQEARLKEPNVTLDEQDILKMTSPDEGLISFLVKNFGNPDRTYK 358

RESULT 3  
 ID P70040 PRELIMINARY; PRT; 382 AA.  
 AC P70040;  
 DT 01-FEB-1997 (Tremblrel. 02, Created)  
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)  
 DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)  
 DE 5' NUCLEASE XFENIA.  
 GN FEN1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Bidikova M., Chi E., Wu B., Kim K., Carroll D.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RA Li J.-L., Cox L.S.;  
 RT "Cloning and investigation of Xenopus Fen1: developmental expression  
 and function in DNA replication."  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U64565; AAB06176.1; -  
 DR EMBL: AF065397; AAD02814.1; -  
 DR HSSP: Q58839; 1A76.  
 DR InterPro: IPR002421; 5\_3\_exonuclease.  
 DR InterPro: IPR000513; Exo\_N\_I.  
 DR InterPro: IPR003584; HHN\_2.  
 DR InterPro: IPR001532; XPG\_I.

DR Pfam: PF01367; 5\_3-exonuclease; 1.  
 DR Pfam: PF00867; XPG\_I; 1.  
 DR Pfam: PF00752; XPG\_N; 1.  
 DR PRINTS: PR00853; XPGADUPER.  
 DR SMART: SM00279; HHH2; 1.  
 DR SMART: SM00484; XPGI; 1.  
 DR SMART: SM00485; XPGN; 1.  
 DR Endonuclease.  
 SW SEQUENCE 382 AA: 42668 MW: 9810B0EAD158D57 CRC64;

Query Match 55.1%; Score 1069; DB 13; Length 382;  
 Best Local Similarity 55.7%; Pred. No. 2.7e-68;  
 Matches 210; Conservative 63; Mismatches 96; Indels 6; Gaps 4;

QY 1 MGIGKLTLLADNAPRAKKEQFESEYFGKRIAVDASMSIYQFLVVGRTGMEITLNEAGE 60  
 DB 1 MGHIGLAKLADYAPAIKEHDIKSYFGKRVAVDASMSIYQFLVAVRODG-NITLONEGE 59  
 QY 61 VTSHLGMEFRTIRLVEHGIRPVYFDGKPPDMKKOELAKRSKRDPAITKDLTEAVESGD 120  
 DB 60 TTHLMGMEFRTIRLVEHGIRPVYFDGKPPDMKKOELAKRSKRDPAITKDLTEAVESGD 119  
 QY 121 KDAIELSKRTYKVTROHNECKRLRLMGVPEVAPSEAEACALCINDKRVAVASED 180  
 DB 120 VENIEKFTRLVLRVYKOHNECKRLRLMGVPEVAPSEAEACALCINDKRVAVASED 179  
 QY 181 MSLTPGAPRFLRHLDSPSSKRIPVMEFPAVLELEITMDQFIDLCILGCDYCSIK 240  
 DB 180 MDALFTGTPVLRHLTASEAKKLPIQEFHLNRVIOIDIGITHEQVDCILIGSDYCETIR 239  
 QY 241 GIGGQALKLITROHSIESILENLKDRYQIPEDMPYOEARLKEKPNVT-LDIPELKWT 299  
 DB 240 GIGPKRAIDLIRQHKITIDEITIDILKRYVPENNLHKAHLFLEPEVDTIDITELKWI 299  
 QY 300 APDEEGLISFLVNDGNEFDRYTKAIEKIKSAKNSGGRSEFPPKTAATSAFLKRET 359  
 DB 300 EPDEGLVAFMCGEKFSEDRIRNGAKKLAKNQGSTGRLDFFKTVGSVSS-TRKEA 358  
 QY 360 SDRTSRAANKTKKAG 376  
 DB 359 E--SKGSARKKAKTGG 372

## RESULT 4

ID 057351 PRELIMINARY; PRT; 382 AA.  
 AC 057351;  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE FLAP ENDONUCLEASE 1.  
 GN FEN1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98204872; PubMed=9535864;  
 RA Kim K., Blaise S., Matsumoto Y.;  
 RT "Involvement of flap endonuclease 1 in base excision DNA repair."  
 RL J. Biol. Chem. 273:8842-8848(1998).  
 DR EMBL: AF036327; AAB88707.1; -  
 DR HSSP: Q58839; 1A76.  
 DR InterPro: IPR002421; 5\_3-exonuclease.  
 DR InterPro: IPR000513; EXO\_N\_I.  
 DR InterPro: IPR003584; HHH\_2.  
 DR InterPro: IPR001532; XPG\_I.  
 DR Pfam: PF01367; 5\_3-exonuclease; 1.  
 DR Pfam: PF00867; XPG\_I; 1.  
 DR Pfam: PF00752; XPG\_N; 1.  
 DR PRINTS: PR00853; XPGADUPER.

DR SMART: SM00279; HHH2; 1.  
 DR SMART: SM00484; XPGI; 1.  
 DR SMART: SM00485; XPGN; 1.  
 DR Endonuclease.  
 SW SEQUENCE 382 AA: 42682 MW: 3A911F83DC585A59 CRC64;

Query Match 55.1%; Score 1068; DB 13; Length 382;  
 Best Local Similarity 55.4%; Pred. No. 3.2e-68;  
 Matches 209; Conservative 64; Mismatches 98; Indels 6; Gaps 4;

QY 1 MGIGKLTLLADNAPRAKKEQFESEYFGKRIAVDASMSIYQFLVVGRTGMEITLNEAGE 60  
 DB 1 MGHIGLAKLADYAPAIKEHDIKSYFGKRVAVDASMSIYQFLVAVRODG-NITLONEGE 59  
 QY 61 VTSHLGMEFRTIRLVEHGIRPVYFDGKPPDMKKOELAKRSKRDPAITKDLTEAVESGD 120  
 DB 60 TTHLMGMEFRTIRLVEHGIRPVYFDGKPPDMKKOELAKRSKRDPAITKDLTEAVESGD 119  
 QY 121 KDAIELSKRTYKVTROHNECKRLRLMGVPEVAPSEAEACALCINDKRVAVASED 180  
 DB 120 VENIEKFTRLVLRVYKOHNECKRLRLMGVPEVAPSEAEACALCINDKRVAVASED 179  
 QY 181 MSLTPGAPRFLRHLDSPSSKRIPVMEFPAVLELEITMDQFIDLCILGCDYCSIK 240  
 DB 180 MDALFTGTPVLRHLTASEAKKLPIQEFHLNRVIOIDIGITHEQVDCILIGSDYCETIR 239  
 QY 241 GIGGQALKLITROHSIESILENLKDRYQIPEDMPYOEARLKEKPNVT-LDIPELKWT 299  
 DB 240 GIGPKRAIDLIRQHKITIDEITIDILKRYVPENNLHKAHLFLEPEVDTIDITELKWI 299  
 QY 300 APDEEGLISFLVNDGNEFDRYTKAIEKIKSAKNSGGRSEFPPKTAATSAFLKRET 359  
 DB 300 EPDEGLVAFMCGEKFSEDRIRNGAKKLAKNQGSTGRLDFFKTVGSVSS-TRKEA 358  
 QY 360 SDRTSRAANKTKKAG 376  
 DB 359 E--SKGSARKKAKTGG 372

## RESULT 5

ID P70054 PRELIMINARY; PRT; 382 AA.  
 AC P70054;  
 DT 01-FEB-1997 (TREMblrel. 02, Created)  
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE XEN1B.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HEAD;  
 RX MEDLINE=99069415; PubMed=9852084;  
 RA Bibikova M., Wu B., Chi E., Kim K.H., Trautman J.K., Carroll D.;  
 RT "Characterization of FEN-1 from Xenopus laevis. cDNA cloning and role  
 in DNA metabolism."  
 RL J. Biol. Chem. 273:34222-34229(1998).  
 DR EMBL: U68141; AAB08478.1; -  
 DR HSSP: Q58839; 1A76.  
 DR InterPro: IPR002421; 5\_3-exonuclease.  
 DR InterPro: IPR000513; EXO\_N\_I.  
 DR InterPro: IPR001532; XPG\_I.  
 DR InterPro: IPR003584; HHH\_2.  
 DR Pfam: PF01367; 5\_3-exonuclease; 1.  
 DR Pfam: PF00867; XPG\_I; 1.  
 DR Pfam: PF00752; XPG\_N; 1.  
 DR PRINTS: PR00853; XPGADUPER.  
 DR SMART: SM00279; HHH2; 1.  
 DR SMART: SM00484; XPGI; 1.  
 DR SMART: SM00485; XPGN; 1.

SQ SEQUENCE 382 AA: 42865 MW: 1C648936A232D460 CRC64:  
 Query Match 55.1%; Score 1068; DB 13; Length 382;  
 Best Local Similarity 56.2%; Pred. No. 3, 2e-68;  
 Matches 212; Conservative 60; Mismatches 99; Indels 6; Gaps 4;  
 QY 1 MGIGLTKLADNAPKAMEKOFESYFGKRIAVDASMSIYQFLIYVGRGTMETLNEAGE 60  
 DB 1 MGIGHGLAKLADVAAPAKREHDIKSYFGKRVAVDASMCYIOFLIYVRODG-NMLONEGE 59  
 QY 61 VTSHLOGMFKRTIRLLEAGIKPVYVDGKPPDMKKOELAKRYSKDDATKDLTEAVEVD 120  
 DB 60 TTSHLMGMFKRTIRLLEAGIKPVYVDGKPPDMKSGELAKRSEKAEKLEAEFEAGE 119  
 QY 121 KDAIEKLSKRTVYVTRONHNDCKRLRLMGVYVPEAPSEAECAALCINDKVFAVASED 180  
 DB 120 VENIEKRNKRLVYVTRONHNDCKRLRLMGVYVPEAPSEAECAALCINDKVFAVASED 179  
 QY 181 MDLTFGAPRLRLHMDPSSKRIPYMEFDVAKVLEELTMDQFLDLCILGCDYCSIR 240  
 DB 180 MDALTFEGTPLYLRLHMLTASEAKRLPIQEFHLNRFQDIGINHQFVLDLCILGSDYCETIR 239  
 QY 241 GIGGOTAKLIRHOGSIESILENLKDRYOIPEDMPYOEARLREPNV-TLDTPELKW 299  
 DB 240 GIGPKRAIDLIRHOKTIEELIDIDLKTPIDPNMLHKEARQLEPEYIDADITELKW 299  
 QY 300 APDEGLISFLVNDGNGFEDRVTKAIEKISAKNKSOGRLSEFPKPTAT-SAPLKRKE 359  
 DB 300 EPDEGLIVAFMGCEKOFSEDRIRNGAKRLAKNKGSTQGRLDPEFKVYGSIS-TKRKEV 358  
 QY 360 SDKTSKAAANKTKRAGK 376  
 DB 359 E--SKGSTRKSKTGG 372  
 RESULT 6  
 ID 090YB0 PRELIMINARY: PRT: 381 AA.  
 AC 090YB0;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE FEN-1 NUCLEASE.  
 GN FEN1.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Matsuzaki Y., Adachi N., Koyama H.;  
 RT "The FEN-1 nuclease is not essential for DNA replication, but is  
 RT required for maintenance of genomic integrity and base excision repair  
 RT in vertebrate cells."  
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AB058602; BAB6507.1; -  
 SQ SEQUENCE 381 AA: 43054 MW: 71432ECBCB4E67AA CRC64:  
 Query Match 54.8%; Score 1063.5; DB 13; Length 381;  
 Best Local Similarity 54.2%; Pred. No. 6, 7e-68;  
 Matches 206; Conservative 67; Mismatches 104; Indels 3; Gaps 3;  
 QY 1 MGIGLTKLADNAPKAMEKOFESYFGKRIAVDASMSIYQFLIYVGRGTMETLNEAGE 60  
 DB 1 MGIGHGLAKLADVAAPAKREHDIKSYFGKRVAVDASMCYIOFLIYVRODG-NMLONEGE 59  
 QY 61 VTSHLOGMFKRTIRLLEAGIKPVYVDGKPPDMKKOELAKRYSKDDATKDLTEAVEVD 120  
 DB 60 TTSHLMGMFKRTIRLLEAGIKPVYVDGKPPDMKSGELAKRSEKAEKLEAEFEAGE 119  
 QY 121 KDAIEKLSKRTVYVTRONHNDCKRLRLMGVYVPEAPSEAECAALCINDKVFAVASED 180

DB 120 ENIEKFSKRLVYVTRONHNDCKRLRLMGVYVPEAPSEAECAALCINDKVFAVASED 179  
 QY 181 MDLTFGAPRLRLHMDPSSKRIPYMEFDVAKVLEELTMDQFLDLCILGCDYCSIR 240  
 DB 180 MDLTFGSPVIRKHLRLTASAKRLPIQEFHLNRLVODLITREOFVLDLCILGSDYCSIR 239  
 QY 241 GIGGOTAKLIRHOGSIESILENLKDRYOIPEDMPYOEARLREPNV-TLDTPELKW 299  
 DB 240 GIGPKRAVELIRHOKTIEELIDIDLKTPIDPNMLHKEARQLEPEYIDADITELKW 299  
 QY 300 APDEGLISFLVNDGNGFEDRVTKAIEKISAKNKSOGRLSEFPKPTAT-TSAPLKRKE 358  
 DB 300 EPDEGLIVAFMGCEKOFSEDRIRNGAKRLAKNKGSTQGRLDPEFKVYGSIS-TKRKEV 359  
 QY 359 TSDKTSKAAANKTKRAGCK 378  
 DB 360 TKGSAKKAKTNSATKAKFK 379  
 RESULT 7  
 ID 091Z50 PRELIMINARY: PRT: 380 AA.  
 AC 091Z50;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE SIMILAR TO FLAG STRUCTURE-SPECIFIC ENDONUCLEASE.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: BC010203; AAH10203.1; -  
 KW Endonuclease  
 SQ SEQUENCE 380 AA: 42623 MW: 1BE903288B4520D CRC64:  
 Query Match 53.8%; Score 1043; DB 11; Length 380;  
 Best Local Similarity 53.8%; Pred. No. 1, 9e-66;  
 Matches 205; Conservative 69; Mismatches 99; Indels 8; Gaps 5;  
 QY 1 MGIGLTKLADNAPKAMEKOFESYFGKRIAVDASMSIYQFLIYVGRGTMETLNEAGE 60  
 DB 1 MGIGHGLAKLADVAAPAKREHDIKSYFGKRVAVDASMCYIOFLIYVRODG-NMLONEGE 59  
 QY 61 VTSHLOGMFKRTIRLLEAGIKPVYVDGKPPDMKKOELAKRYSKDDATKDLTEAVEVD 120  
 DB 60 TTSHLMGMFKRTIRLLEAGIKPVYVDGKPPDMKSGELAKRSEKAEKLEAEFEAGE 119  
 QY 121 KDAIEKLSKRTVYVTRONHNDCKRLRLMGVYVPEAPSEAECAALCINDKVFAVASED 180  
 DB 120 EEEVERFTKRLVYVTRONHNDCKRLRLMGVYVPEAPSEAECAALCINDKVFAVASED 179  
 QY 181 MDLTFGSPVIRKHLRLTASAKRLPIQEFHLNRLVODLITREOFVLDLCILGSDYCSIR 240  
 DB 180 MDLTFGSPVIRKHLRLTASAKRLPIQEFHLNRLVODLITREOFVLDLCILGSDYCSIR 239  
 QY 241 GIGGOTAKLIRHOGSIESILENLKDRYOIPEDMPYOEARLREPNV-TLDTPELKW 298  
 DB 240 GIGPKRAVELIRHOKTIEELIDIDLKTPIDPNMLHKEARQLEPEYIDADITELKW 298  
 QY 299 TAPDEGLISFLVNDGNGFEDRVTKAIEKISAKNKSOGRLSEFPKPTAT-TSAPLKRKE 358  
 DB 299 SEFNEBELVAFMGCEKOFSEDRIRNGAKRLAKNKGSTQGRLDPEFKVYGSIS-TKRKE 357  
 QY 359 TSDKTSKAAANKTKRAGCK 379  
 DB 358 PE--PKGPAKKAKTGGAGK 375  
 RESULT 8

Q9JHW7 PRELIMINARY; PRT; 380 AA.  
AC Q9JHW7;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE FLAG STRUCTURE-SPECIFIC ENDONUCLEASE.  
GN FEN-1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chen D., Gao G., Yang S., Li M., Chen J.;  
RT "Cloning and characterization of a rat DNA structure-specific  
RT endonuclease (Fen-1)."; the EMBL/GenBank/DBJ databases.  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF281018; AAF81265.1; -.  
DR HSSP; O58839; 1A76.  
DR InterPro; IPR002421; 5\_3\_exonuclease.  
DR InterPro; IPR000513; Exo\_N.1.  
DR InterPro; IPR001532; XPG\_1.  
DR Pfam; PF01367; 5\_3\_exonuclease; 1.  
DR Pfam; PF00867; XPG\_1; 1.  
DR Pfam; PF00752; XPG\_1; 1.  
DR PRINTS; PR00853; XPGRADSUPER.  
DR SMART; SM00279; HhH2; 1.  
DR SMART; SM00484; XPG1; 1.  
DR SMART; SM00485; XPGN; 1.  
DR PROSITE; PS00842; XPG\_2; 1.  
KW Endonuclease.  
SQ  
SEQUENCE 380 AA; 42622 MW; 61698CE1F182136 CRC64;  
Query Match 53.6%; Score 1040; DB 11; Length 380;  
Best Local Similarity 52.6%; Pred. No. 3.1e-66;  
Matches 200; Conservative 73; Mismatches 101; Indels 6; Gaps 4;  
QY 1 MGKIGLTKLADNAPKANKKQKFESEYFGKRIAVDASMSIYOFLLIYGRGMEITLTNEAGE 60  
1 MGKIGLTKLADNAPKANKKQKFESEYFGKRIAVDASMSIYOFLLIYGRGMEITLTNEAGE 59  
DB 1 MGKIGLTKLADNAPKANKKQKFESEYFGKRIAVDASMSIYOFLLIYGRGMEITLTNEAGE 59  
QY 61 VTSHLQGMFNRTIRLEAGIRPVYFDGKPPDMKKQELARYSKRDATKDLTEAVEVD 120  
1 VTSHLQGMFNRTIRLEAGIRPVYFDGKPPDMKKQELARYSKRDATKDLTEAVEVD 120  
DB 60 TTSHLQGMFNRTIRLEAGIRPVYFDGKPPDMKKQELARYSKRDATKDLTEAVEVD 119  
QY 121 KDAIEKLSKRTVAVRQNHDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180  
1 KDAIEKLSKRTVAVRQNHDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180  
DB 120 EEEVEKFTKLVKTVKQNHDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 179  
QY 181 MDLTFGAPRFLRHLMDPSSKIPVMEFDVAKVLELELTMDQFIDLCILGCDYCDISIK 240  
1 MDLTFGAPRFLRHLMDPSSKIPVMEFDVAKVLELELTMDQFIDLCILGCDYCDISIK 240  
DB 180 MDLTFGAPRFLRHLMDPSSKIPVMEFDVAKVLELELTMDQFIDLCILGCDYCDISIK 239  
QY 241 GIGGOTALKLIPHOGSISILELNKDRQIPEDMVPOEARLFEKPNVT-LDIPELKWT 299  
1 GIGGOTALKLIPHOGSISILELNKDRQIPEDMVPOEARLFEKPNVT-LDIPELKWT 299  
DB 240 GIGGOTALKLIPHOGSISILELNKDRQIPEDMVPOEARLFEKPNVT-LDIPELKWT 299  
QY 300 APDEGLISFLVKNQNGFNEDRVYKAIEKISAKKSSQGRLESEFFKPTATTSAPLKRET 359  
1 APDEGLISFLVKNQNGFNEDRVYKAIEKISAKKSSQGRLESEFFKPTATTSAPLKRET 359  
DB 300 APDEGLISFLVKNQNGFNEDRVYKAIEKISAKKSSQGRLESEFFKPTATTSAPLKRET 358  
QY 360 SDKTSKAANKTKAGGKKK 379  
1 SDKTSKAANKTKAGGKKK 379  
DB 359 E---PKGPARKKAKGTGACK 375

RESULT 9  
Q9N3T2 PRELIMINARY; PRT; 382 AA.  
AC Q9N3T2;  
DT 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE HYPOTHETICAL 42.5 KDA PROTEIN.  
GN Y4766A.8.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-BRISTOL N2;  
RC MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA STRAIN-BRISTOL N2;  
RC MEDLINE=99069613; PubMed=9851916;  
RA Minx P., Graves T., Hawrysko C.;  
RT "The sequence of C. elegans cosmid Y4766A.";  
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RP SEQUENCE FROM N.A.  
RA STRAIN-BRISTOL N2;  
RC WATERSTON R.;  
RT "Direct Submission.";  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC024791; AAF60653.1; -.  
DR HSSP; O58839; 1A76.  
DR InterPro; IPR002421; 5\_3\_exonuclease.  
DR InterPro; IPR000513; Exo\_N.1.  
DR InterPro; IPR001532; XPG\_1.  
DR Pfam; PF01367; 5\_3\_exonuclease; 1.  
DR Pfam; PF00867; XPG\_1; 1.  
DR Pfam; PF00752; XPG\_1; 1.  
DR PRINTS; PR00853; XPGRADSUPER.  
DR SMART; SM00279; HhH2; 1.  
DR SMART; SM00484; XPG1; 1.  
DR SMART; SM00485; XPGN; 1.  
KW Hypothetical protein.  
SQ  
SEQUENCE 382 AA; 42549 MW; EFD60BA0707321EB CRC64;  
Query Match 53.5%; Score 1036.5; DB 5; Length 382;  
Best Local Similarity 52.3%; Pred. No. 5.6e-66;  
Matches 201; Conservative 73; Mismatches 99; Indels 11; Gaps 4;  
QY 1 MGKIGLTKLADNAPKANKKQKFESEYFGKRIAVDASMSIYOFLLIYGRGMEITLTNEAGE 60  
1 MGKIGLTKLADNAPKANKKQKFESEYFGKRIAVDASMSIYOFLLIYGRGMEITLTNEAGE 59  
DB 1 MGKIGLTKLADNAPKANKKQKFESEYFGKRIAVDASMSIYOFLLIYGRGMEITLTNEAGE 59  
QY 61 VTSHLQGMFNRTIRLEAGIRPVYFDGKPPDMKKQELARYSKRDATKDLTEAVEVD 120  
1 VTSHLQGMFNRTIRLEAGIRPVYFDGKPPDMKKQELARYSKRDATKDLTEAVEVD 120  
DB 60 TTSHLQGMFNRTIRLEAGIRPVYFDGKPPDMKKQELARYSKRDATKDLTEAVEVD 119  
QY 121 KDAIEKLSKRTVAVRQNHDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180  
1 KDAIEKLSKRTVAVRQNHDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180  
DB 120 VEAEEKFTKLVKTVKQNHDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 179  
QY 181 MDLTFGAPRFLRHLMDPSSKIPVMEFDVAKVLELELTMDQFIDLCILGCDYCDISIK 240  
1 MDLTFGAPRFLRHLMDPSSKIPVMEFDVAKVLELELTMDQFIDLCILGCDYCDISIK 240  
DB 180 MDLTFGAPRFLRHLMDPSSKIPVMEFDVAKVLELELTMDQFIDLCILGCDYCDISIK 239  
QY 241 GIGGOTALKLIPHOGSISILELNKDRQIPEDMVPOEARLFEKPNVT-LDIPELKWT 299  
1 GIGGOTALKLIPHOGSISILELNKDRQIPEDMVPOEARLFEKPNVT-LDIPELKWT 299  
DB 240 GIGGOTALKLIPHOGSISILELNKDRQIPEDMVPOEARLFEKPNVT-LDIPELKWT 299  
QY 300 APDEGLISFLVKNQNGFNEDRVYKAIEKISAKKSSQGRLESEFFKPTATTSAPLKRET 355  
1 APDEGLISFLVKNQNGFNEDRVYKAIEKISAKKSSQGRLESEFFKPTATTSAPLKRET 355  
DB 300 EADVEGVYDFGLGKFNEDRVYKAIEKISAKKSSQGRLESEFFKPTATTSAPLKRET 359  
QY 356 RRTSDKTSKAANKTKAGGKKK 379

Db 360 KAEAEKAKKGA-----KKSGPPK 378

```
RESULT 10
ID 096830 PRELIMINARY: PRT: 378 AA.
AC 096830:
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE EG:EG0003.3 OR CG8648.
GN Drosophila melanogaster (fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Adams M.D., Gelinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Goddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hastin N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Murphey L., Harris D., Barrell B.;
RT "Sequencing the distal X chromosome of Drosophila melanogaster."
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Benos P.;
RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AEO03805; AAF57944.1; -.
DR EMBL: AL031863; CAA21320.1; -.
DR HSSP: Q58839; 1A76.
DR FlyBase: FBgn0025832; Feni.
DR InterPro: IPR000513; Exo_N_I.
DR InterPro: IPR000513; Exo_N_I.
```

```
RESULT 11
ID 096830 PRELIMINARY: PRT: 385 AA.
AC 096830:
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE EG:EG0003.3 OR CG8648.
GN Drosophila melanogaster (fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Adams M.D., Gelinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Goddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hastin N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Murphey L., Harris D., Barrell B.;
RT "Sequencing the distal X chromosome of Drosophila melanogaster."
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Benos P.;
RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AEO03805; AAF57944.1; -.
DR EMBL: AL031863; CAA21320.1; -.
DR HSSP: Q58839; 1A76.
DR FlyBase: FBgn0025832; Feni.
DR InterPro: IPR000513; Exo_N_I.
DR InterPro: IPR000513; Exo_N_I.
```







```
Db 132 IEDAKKLLQLMGIPYVAPSEGEAQAAYMAGGDVYASASODYDSILFGTPRLVNRNLTIT 191
QY 199 SSKKIPWMEFDV-----AKVEELELTMDQFIDLCILGCDY-CDSIKIGIGOTAL 248
Db 192 GKRMPGKDIYVEIKPELIVLEVELEKELITREKLELAIWGTDYNEGIKIGIPKKAL 251
QY 249 KLIRHGSIESILENLNKD--RYQIPEDWPYQEARRLFKEPNVTLIDIPELKWTAPDEEG 305
Db 252 EIVK-----YKDBLAKFQROSDVDLYAIKEFFLNPTTDDY-SLKMRPDEEG 299
QY 306 LISPLVNDNGFNEDRYTKALEIKSAKNSSQGRLESEF--KP 346
Db 300 IIRFLCDEHDSERERKNGLELKLKAIKAKOSTLESWPIKKP 343
```

Search completed: November 5, 2002, 16:29:04  
Job time : 30 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: November 5, 2002, 09:27.12 ; Search time 1750.69 Seconds  
(without alignments)  
16507.481 Million cell updates/sec

Title: US-09-805-311-5

Perfect score: 1381

Sequence: 1 cgaccacgcgtccgcgcac.....ttgaaaaaaaaaaaaaaaaa 1381

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_bhg:\*  
3: gb\_in:\*  
4: gb\_com:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mta:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hicg\_hum:\*  
31: em\_hicg\_inv:\*  
32: em\_hicg\_other:\*  
33: em\_hicgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
------------	-------	-------------	--------	----	----	-------------

1	1381	100.0	1381	6	ARI52405	ARI52405 Sequence
2	1350.2	97.8	1463	6	ARI52403	ARI52403 Sequence
3	1348.6	97.7	1541	6	ARI52404	ARI52404 Sequence
4	1334.2	96.6	1478	6	ARI52406	ARI52406 Sequence
5	924	66.9	1354	8	AB021666	AB021666 Oryza sat
6	326.8	23.7	1469	5	AB021666	AB021666 Oryza sat
7	313.8	22.7	1461	5	AB021666	AB021666 Oryza sat
8	307	22.2	1357	5	AF065397	AF065397 Xenopus lae
9	305.4	22.1	1149	5	AF065397	AF065397 Xenopus lae
10	305.4	22.1	1149	5	AF065397	AF065397 Xenopus lae
11	304.2	22.1	16086	2	BC010203	BC010203 Mus muscu
12	304.2	22.0	1757	9	HSRAD2	HSRAD2
13	304.2	22.0	2031	9	BC000323	BC000323 Homo sapi
14	302.8	21.9	138627	2	AP000591	AP000591 Homo sapi
15	302.8	21.9	169053	2	AP000591	AP000591 Homo sapi
16	302.8	21.9	185035	9	AC004770	AC004770 Homo sapi
17	302.8	21.9	196080	9	AC004770	AC004770 Homo sapi
18	302.6	21.9	1144	9	HUMEN1A	HUMEN1A
19	299	21.7	1146	5	AB058602	AB058602 Gallus ga
20	292.4	21.2	39563	9	HSU73629	HSU73629 Human chro
21	288.2	20.9	2192	10	AF281018	AF281018 Rattus no
22	278.8	20.3	2033	10	MUSEF1X	MUSEF1X
23	278.2	20.1	6361	10	AF014962	AF014962 Mus muscu
24	276.6	19.9	2114	8	SCYKL113C	SCYKL113C
25	274.6	19.9	10660	8	S93804	S93804 BAF1/ABF1/Y
26	260.4	18.9	151865	2	AL356236	AL356236 Homo sapi
27	260.4	18.9	179461	2	AC019177	AC019177 Homo sapi
28	259.4	18.8	2299	8	SPRAD2	SPRAD2
29	259.4	18.8	28491	8	SPAC366	SPAC366
30	258.8	18.7	173256	2	AC073301	AC073301 Homo sapi
31	258.8	18.7	178806	2	AC013557	AC013557 Homo sapi
32	252.6	18.3	37688	2	AL513222	AL513222 Drosophila
33	252.6	18.3	37688	2	DMCEG0003	DMCEG0003
34	252.6	18.3	95880	2	AC020003	AC020003
35	252.6	18.3	15985	3	AC004641	AC004641 Drosophila
36	252.6	18.3	160338	3	AC099021	AC099021 Drosophila
37	252.6	18.3	270306	3	AE003805	AE003805 Drosophila
38	241.8	17.5	194322	2	AC024791	AC024791 Caenorhab
39	241.8	17.5	319857	2	AC006782	AC006782 Caenorhab
40	237	17.2	2783	3	AF093702	AF093702 Plasmodiu
41	237	17.2	234112	3	PFMAL4P2	PFMAL4P2
42	236.6	17.1	1953	3	AF278764	AF278764 Plasmodiu
43	236.4	17.1	194439	3	CNS0786A	CNS0786A
44	233.6	16.9	120848	2	AC007801	AC007801 Drosophila
45	159.6	11.6	1023	6	AR062093	AR062093 Sequence

#### ALIGNMENTS

RESULT 1  
LOCUS ARI52405 ARI52405 1381 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 5 from patent US 6232527.  
ACCESSION ARI52405  
VERSION ARI52405.1 GI:15118455  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1381)  
AUTHORS Mahajan,P.B.  
TITLE Maize Rad3/FEN-1 orthologues and uses thereof  
JOURNAL Patent: US 6232527-A 5 15-MAY-2001;  
FEATURES  
source Location/Qualifiers  
1..1381  
BASE COUNT 441 a 269 c 346 g 325 t  
ORIGIN

Query Match 100.0%; Score 1381; DB 6; Length 1381;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGACGACGCGCGGACGACGCGCGCGGACGAGATGAGGCGTCAAGAGGTTTGACGAA	60
Db	1	CGACCGACGCGCGCGGACGCGCGCGGACGAGAGATGAGGCGTCAAGAGGTTTGACGAA	60
QY	61	CTGTCTGGGAGACATGCGGCCCAAGCGGATGAGAGGACGAGAGTTGAGAGACTTGTGGC	120
Db	61	CTGTCTGGGAGACATGCGGCCCAAGCGGATGAGAGGACGAGAGTTGAGAGACTTGTGGC	120
QY	121	CGCAAAATGCGCGTGGAGCGCGACGATGACCATATACAGTTCTGTATGTGTGTGAAAG	180
Db	121	CGCAAAATGCGCGTGGAGCGCGACGATGACCATATACAGTTCTGTATGTGTGTGAAAG	180
QY	181	ACAGGCGATGAAACCTCACAAATGAAAGCTGGTGAAGTCACTATCTATTTGCAAGAAATG	240
Db	181	ACAGGCGATGAAACCTCACAAATGAAAGCTGGTGAAGTCACTATCTATTTGCAAGAAATG	240
QY	241	TTCAACCGGACAAATAGATTCTGGAAGGGGAGTCAAGCCAGTTATGTATTGTATGGC	300
Db	241	TTCAACCGGACAAATAGATTCTGGAAGGGGAGTCAAGCCAGTTATGTATTGTATGGC	300
QY	301	AACCTCTCGATATGAGAAAGAAACAGAGCTTGTAAABATCTCAAAABAGATGATGA	360
Db	301	AACCTCTCGATATGAGAAAGAAACAGAGCTTGTAAABATCTCAAAABAGATGATGA	360
QY	361	ACCAAAGATCTGACTGAGCGAGTGAAGTGAAGATGAAAGTGCATGTGAAAAATTGAC	420
Db	361	ACCAAAGATCTGACTGAGCGAGTGAAGTGAAGATGAAAGTGCATGTGAAAAATTGAC	420
QY	421	AACGAGACTGTAAAGSTACAAAGCAACAAACGAAGATTTTAAACGGCTATTAGACTT	480
Db	421	AACGAGACTGTAAAGSTACAAAGCAACAAACGAAGATTTTAAACGGCTATTAGACTT	480
QY	481	ATGGGGTCTCGATTGTAGAGCGACCTTCGAGCAGAGAGAGANTGTGAGCCCTTTC	540
Db	481	ATGGGGTCTCGATTGTAGAGCGACCTTCGAGCAGAGAGAGANTGTGAGCCCTTTC	540
QY	541	ATTAACGATTAAGGTGTGGCTGTCTCAGAAATATGAGACTCCCTACTTTGGGGCT	600
Db	541	ATTAACGATTAAGGTGTGGCTGTCTCAGAAATATGAGACTCCCTACTTTGGGGCT	600
QY	601	CGAAGTTCCTTCTGCTATTATGATCAAGTTCCAGAAATATACCTGTGATGGAATTT	660
Db	601	CGAAGTTCCTTCTGCTATTATGATCAAGTTCCAGAAATATACCTGTGATGGAATTT	660
QY	661	GATCTTGGCCAGGTTTGGAGAGGACTTGAACCTACCATGAGCAAGTTCAATTGATTGTGC	720
Db	661	GATCTTGGCCAGGTTTGGAGAGGACTTGAACCTACCATGAGCAAGTTCAATTGATTGTGC	720
QY	721	ATTCCTGTGAGATGTGATATTGTGTGATGATCAAGAGATGAGGGGGCAAAAGCTCTG	780
Db	721	ATTCCTGTGAGATGTGATATTGTGTGATGATCAAGAGATGAGGGGGCAAAAGCTCTG	780
QY	781	AAACTATTCTGACACATAGGGTCCATGAAAGATCTTGGAGANTCTTATTAAGACAGA	840
Db	781	AAACTATTCTGACACATAGGGTCCATGAAAGATCTTGGAGANTCTTATTAAGACAGA	840
QY	841	TATCAAAATTCCTGAGAGACTGCGCTTACCAAGAGCTCGACGCTGTCTTCAAGAGCTAAT	900
Db	841	TATCAAAATTCCTGAGAGACTGCGCTTACCAAGAGCTCGACGCTGTCTTCAAGAGCTAAT	900
QY	901	GTCACATTGTGAATTCCTGAGACTAAATGACATGCACTGATGAGGAGGGCTCATTAAT	960
Db	901	GTCACATTGTGAATTCCTGAGACTAAATGACATGCACTGATGAGGAGGGCTCATTAAT	960
QY	961	TTCTCTGTTAAAGATATAGTTTTCACGAAGATCGGGGTACAAAGGCGCATRAGAAATATC	1020
Db	961	TTCTCTGTTAAAGATATAGTTTTCACGAAGATCGGGGTACAAAGGCGCATRAGAAATATC	1020
QY	1021	AAATCTGCCAACAATTAATCTGTCCAGAGAACTGAGTCTTTTCAACCAACTGCC	1080
Db	1021	AAATCTGCCAACAATTAATCTGTCCAGAGAACTGAGTCTTTTCAACCAACTGCC	1080

QY	1081	ACCCATACACCCCGGTAAAGCGAAGGAGACTTCGGATTAACCAAGCAAGGAGCTGG	1140
Db	1081	ACGCATCAGACACCGGTAAAGCGAAGGAGACTTCGGATTAACCAAGCAAGGAGCTGG	1140
QY	1141	AACCAAGAAAACCAAGGCTGGTGGAAAGAAATAATCTTGATGATGCTGATGATCAACCA	1200
Db	1141	AACCAAGAAAACCAAGGCTGGTGGAAAGAAATAATCTTGATGATGCTGATGATCAACCA	1200
QY	1201	CGACTACGAAGCAGCGGTGGCGCTGATGCCTTAGATTATTTACTCCCTGTTTA	1260
Db	1201	CGACTACGAAGCAGCGGTGGCGCTGATGCCTTAGATTATTTACTCCCTGTTTA	1260
QY	1261	ACTCAGAGCTTTGGTAAAGTTGCGCCATGTTTCAAGCTGGAGTAAGTTAGTTGCTTGG	1320
Db	1261	ACTCAGAGCTTTGGTAAAGTTGCGCCATGTTTCAAGCTGGAGTAAGTTAGTTGCTTGG	1320
QY	1321	AAGGATTTGGTGATCAACATTAACCAACTTTCGCTGTTTGGAAAAA	1380
Db	1321	AAGGATTTGGTGATCAACATTAACCAACTTTCGCTGTTTGGAAAAA	1380
QY	1381	A 1381	
Db	1381	A 1381	
RESULT 2			
LOCUS	ARI52403	1463 bp	DNA linear PAT 08-AUG-2001
DEFINITION	Sequence 1 from patent US 6232527.		
ACCESSION	ARI52403		
VERSION	ARI52403.1	GI:15118453	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1463)		
TITLE	Mahajan, P. B.		
JOURNAL	Maize Rad2/FEN-1 orthologues and uses thereof		
FEATURES	Patent: US 6232527-A 1 15-MAY-2001;		
source	Location/Qualifiers		
	1..1463		
	/organism="unknown"		
BASE COUNT	466 a 292 c 361 g 344 t		
ORIGIN			
Query Match	97.8%;	Score 1350.2;	DB 6; Length 1463;
Best Local Similarity	99.4%;	Pred. No. of	
Matches 135;	Conservative	0; Mismatches	8; Indels 0; Gaps 0;
QY	1	CGAGCCACGCGTCCGGCCAGCGCCGCGCAGAGATGGGATCAAGGCTTTGACGAAA	60
Db	49	CGCGCCCGCCGACCGCCGACAGCGCCGCGCAGAGATGGGATCAAGGCTTTGACGAAA	108
QY	61	CTGCTGGCGGACAAATGCGGCCCAAGGCGATGAAGGAGAGAGTTGAGACTTACCTTCGC	120
Db	109	CTGCTGGCGGACAAATGCGGCCCAAGGCGATGAAGGAGAGAGTTGAGACTTACCTTCGC	168
QY	121	CGCAAAATGCGCCGCGCCAGCGCGCATGAGCATATATCCGATTCGATTGTAATGGGAAG	180
Db	169	CGCAAAATGCGCCGCGCCAGCGCGCATGAGCATATATCCGATTCGATTGTAATGGGAAG	228
QY	181	ACAAGCATGGAATCTCTACAAATGGAAGCTGTGAAGTCACTAGTCAATTTGCAAGAAATG	240
Db	229	ACAAGCATGGAATCTCTACAAATGGAAGCTGTGAAGTCACTAGTCAATTTGCAAGAAATG	288
QY	241	TTTCAACGGGCAATTAAGATTAAGTACGAAAGGGGGAATCAAGCCATTTTATGTTTGAATGC	300
Db	289	TTTCAACGGGCAATTAAGATTAAGTACGAAAGGGGGAATCAAGCCATTTTATGTTTGAATGC	348
QY	301	AAGCTCTCTATATGAAGAAACAAGCTCTCTATAAGATACATCAAAAAGAGATGATGA	360
Db	349	AAGCTCTCTATATGAAGAAACAAGCTCTCTATAAGATACATCAAAAAGAGATGATGA	408

QY 361 ACCAAGATCTGACTGAGCGCAGTAGAGTAGAGATTAAGATCGATTGAAAAATTGAGC 420  
|||||  
Db 409 ACCAAGATCTGACTGAGCGCAGTAGAGTAGAGATTAAGATCGATTGAAAAATTGAGC 468  
QY 421 AAGAGGCTCTAAAGGTCTACAAAGGCAACCAAGAAATGTTAAAGCGCTATTAGACTT 480  
|||||  
Db 469 AAGAGGCTCTAAAGGTCTACAAAGGCAACCAAGAAATGTTAAAGCGCTATTAGACTT 528  
QY 481 ATGGGGGTTCTGTGTAGAGGACCTTGTGAAGCAGAAGCAAGATGTGCAAGCCCTTGGC 540  
|||||  
Db 529 ATGGGGGTTCTGTGTAGAGGACCTTGTGAAGCAGAAGCAAGATGTGCAAGCCCTTGGC 588  
QY 541 ATAAAGCATTAAGTGTTCGCTGTGCTTCAAGATATGACTCCCTTACTTTTGGGCT 600  
|||||  
Db 589 ATAAAGCATTAAGTGTTCGCTGTGCTTCAAGATATGACTCCCTTACTTTTGGGCT 648  
QY 601 CCAAGGTTCTCTGTATTAATGATCCAAATCCAAAGAAATACCTGTGATGAAATTT 660  
|||||  
Db 649 CCAAGGTTCTCTGTATTAATGATCCAAATCCAAAGAAATACCTGTGATGAAATTT 708  
QY 661 GATGTTCCCAAGGTTTGGAGAGCTTGAAGTCCAGTCCAGTCAATGATTTGTC 720  
|||||  
Db 709 GATGTTCCCAAGGTTTGGAGAGCTTGAAGTCCAGTCCAGTCAATGATTTGTC 768  
QY 721 ATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
|||||  
Db 769 ATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 828  
QY 781 AAATTAATTTGTCACATGAGGCTCATAGAAAGCATTTGGAGATCTTAAATTAAGACGA 840  
|||||  
Db 829 AAATTAATTTGTCACATGAGGCTCATAGAAAGCATTTGGAGATCTTAAATTAAGACGA 888  
QY 841 TATCAAAATTCCTGAGAGACTGGCTTACCAAGAGCTCGAGCTTGTCAAGAGGCTTAT 900  
|||||  
Db 889 TATCAAAATTCCTGAGAGACTGGCTTACCAAGAGCTCGAGCTTGTCAAGAGGCTTAT 948  
QY 901 GTGACATGATTAATTCCTAGAGCTAAATGATGATGATGATGATGATGATGATGATG 960  
|||||  
Db 949 GTGACATGATTAATTCCTAGAGCTAAATGATGATGATGATGATGATGATGATGATG 1008  
QY 961 TTCTGTGTAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
|||||  
Db 1009 TTCTGTGTAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1068  
QY 1021 AAATTCGCAAGATTAATTCCTAGAGCTAAATGATGATGATGATGATGATGATGATG 1080  
|||||  
Db 1069 AAATTCGCAAGATTAATTCCTAGAGCTAAATGATGATGATGATGATGATGATGATG 1128  
QY 1081 ACCACATCAGCAGCGCTAAAGCGAAGAGAGACTTGGATTAACAAGCAAGCAGCTGCG 1140  
|||||  
Db 1129 ACCACATCAGCAGCGCTAAAGCGAAGAGAGACTTGGATTAACAAGCAAGCAGCTGCG 1188  
QY 1141 AACAGAAACAAAGGCTGTGTAAGAAAGAAATATCTTGATGCTGATGATGATGATGATG 1200  
|||||  
Db 1189 AACAGAAACAAAGGCTGTGTAAGAAAGAAATATCTTGATGCTGATGATGATGATGATG 1248  
QY 1201 CGACTACGAAGAGCGGTGCGTATGATGATGATGATGATGATGATGATGATGATG 1260  
|||||  
Db 1249 CGACTACGAAGAGCGGTGCGTATGATGATGATGATGATGATGATGATGATGATG 1308  
QY 1261 ACCTAGAGCTTGTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
|||||  
Db 1309 ACCTAGAGCTTGTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1368  
QY 1321 AAGAGATTGTTGATCAAGATTAACAAAGCTTATGCTGTTT 1363  
|||||  
Db 1369 AAGAGATTGTTGATCAAGATTAACAAAGCTTATGCTGTTT 1411  
RESULT 3  
ARI52404  
LOCUS ARI52404 1541 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 3 from patent US 6232527.

ACCESSION ARI52404  
VERSION ARI52404.1 GI:15118454  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 1541)  
TITLE Maize Rad2/FEN-1 orthologues and uses thereof  
JOURNAL Patent: US 6232527-A 3 15-MAY-2001;  
FEATURES  
source 1. 1541  
BASE COUNT 473 a 308 c 377 g 383 t  
ORIGIN  
Query Match 97.7%; Score 1348.6; DB 6; Length 1541;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1354; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 1 CGACCACGCGTCCGCCACAGCCGCCAGACGAGATGAGCATCAAGGTTTGACGAAA 60  
|||||  
Db 43 CGCCGCCGCCACCCGCCACAGCCGCCAGACGAGATGAGCATCAAGGTTTGACGAAA 102  
QY 61 CTGCTGGCGGACAAATGCGCCCAAGGCGATGAAGGAGAGAGATGAGAGTCTTGCGC 120  
|||||  
Db 103 CTGCTGGCGGACAAATGCGCCCAAGGCGATGAAGGAGAGAGATGAGAGTCTTGCGC 162  
QY 121 CGCAAAATCGCCGTGACGCCAGCATGACATATACAGTTCCTGATTTGATGGAAG 180  
|||||  
Db 163 CGCAAAATCGCCGTGACGCCAGCATGACATATACAGTTCCTGATTTGATGGAAG 222  
QY 181 ACAGGATGAAAGCTCTCAAAATGAGCTGATGATGATGATGATGATGATGATGATG 240  
|||||  
Db 223 ACAGGATGAAAGCTCTCAAAATGAGCTGATGATGATGATGATGATGATGATGATG 282  
QY 241 TTCAACCGCAATTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATG 300  
|||||  
Db 283 TTCAACCGCAATTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATG 342  
QY 301 AAGCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
|||||  
Db 343 AAGCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 402  
QY 361 ACCAAGATCTGACTGAGCGCAGTAGAGTAGAGATTAAGATCGATTGAAAAATTGAGC 420  
|||||  
Db 403 ACCAAGATCTGACTGAGCGCAGTAGAGTAGAGATTAAGATCGATTGAAAAATTGAGC 462  
QY 421 AAGAGACTGTAAAGGTCTCAAGGCAACCAAGAGATTTGAACGGCTATTAGACTT 480  
|||||  
Db 463 AAGAGACTGTAAAGGTCTCAAGGCAACCAAGAGATTTGAACGGCTATTAGACTT 522  
QY 481 ATGGGGGTTCTGTGTAGAGGACCTTGTGAAGCAGAAGCAAGATGTGCAAGCCCTTGGC 540  
|||||  
Db 523 ATGGGGGTTCTGTGTAGAGGACCTTGTGAAGCAGAAGCAAGATGTGCAAGCCCTTGGC 582  
QY 541 ATAAAGCATTAAGTGTTCGCTGTGCTTCAAGATATGACTCCCTTACTTTTGGGCT 600  
|||||  
Db 583 ATAAAGCATTAAGTGTTCGCTGTGCTTCAAGATATGACTCCCTTACTTTTGGGCT 642  
QY 601 CCAAGGTTCTCTGTATTAATGATCCAAATCCAAAGAAATACCTGTGATGAAATTT 660  
|||||  
Db 643 CCAAGGTTCTCTGTATTAATGATCCAAATCCAAAGAAATACCTGTGATGAAATTT 702  
QY 661 GATGTTCCCAAGGTTTGGAGAGCTTGAAGTCCAGTCCAGTCAATGATTTGTC 720  
|||||  
Db 703 GATGTTCCCAAGGTTTGGAGAGCTTGAAGTCCAGTCCAGTCAATGATTTGTC 762  
QY 721 ATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
|||||  
Db 763 ATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 822  
QY 781 AAATTAATTTGTCACATGAGGCTCATAGAAAGCATTTGGAAGATCTTAAATTAAGACGA 840

```
|||||
Db 823 AACTTTTGTGACATGGGTCCATAGAAAGCATCTTGGAGAACTTAAATTAACAGCA 882
Qy 841 TATCAAAATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 900
Db 883 TATCAAAATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 942
Qy 901 GTCAATTTGATATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 960
Db 943 GTCAATTTGATATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 1002
Qy 961 TTCTGTGTAAGAAATTAATGTTTCAACGAAGATGGGTGACAAAGGCTTACAGAGATC 1020
Db 1003 TTCTGTGTAAGAAATTAATGTTTCAACGAAGATGGGTGACAAAGGCTTACAGAGATC 1062
Qy 1021 AATCTGTCAGAAATTAATGTTTCAACGAAGATGGGTGACAAAGGCTTACAGAGATC 1080
Db 1063 AATCTGTCAGAAATTAATGTTTCAACGAAGATGGGTGACAAAGGCTTACAGAGATC 1122
Qy 1081 ACCATTCAGACCGCTTAAACGGAAGAGACTTGGATTAACAAAGGCTTACAGAGATC 1140
Db 1123 ACCATTCAGACCGCTTAAACGGAAGAGACTTGGATTAACAAAGGCTTACAGAGATC 1182
Qy 1141 AACAGAAACAAAGGCTTGGTGAAGAAAGAAATTAATCTTGGATGCTTATGATACACTA 1200
Db 1183 AACAGAAACAAAGGCTTGGTGAAGAAAGAAATTAATCTTGGATGCTTATGATACACTA 1242
Qy 1201 CGACTACGAAGAGCGGTGGCTGATCAGCTTGCCTTATGATTAATCTTCCCTGTTTGA 1260
Db 1243 CGACTACGAAGAGCGGTGGCTGATCAGCTTGCCTTATGATTAATCTTCCCTGTTTGA 1302
Qy 1261 ACTCAGAGCTTGTGTAAGAAATTAATGTTTCAACGAGCTGGGTAGTAGTGTGTGTTG 1320
Db 1303 ACTCAGAGCTTGTGTAAGAAATTAATGTTTCAACGAGCTGGGTAGTAGTGTGTGTTG 1362
Qy 1321 AAGAGATTGGTGTACCAAGTAAACAAATTAATGCTGTTTGT 1363
Db 1363 AAGAGATTGGTGTACCAAGTAAACAAATTAATGCTGTTTGT 1405

RESULT 4
ARI52406 1478 bp DNA linear PAT 08-AUG-2001
LOCUS ARI52406
DEFINITION Sequence 7 from patent US 6232527.
ACCESSION ARI52406
VERSION ARI52406.1 GI:15118456
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1478)
AUTHORS
Mahajan,P.B.
TITLE
Maize Rad2/FEN-1 orthologues and uses thereof
JOURNAL
Patent: US 6232527-A 7 15-MAY-2001;
FEATURES
location/Qualifiers
1..1478
source
BASE COUNT 463 a 302 c 365 g 348 t
ORIGIN
Query Match 96.6%; Score 1334.2; DB 6; Length 1478;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1345; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

|||||
Db 181 CGCAAAATGCGCGTGCAGCGCCAGCATGAGCATCTACGATTCCTGATAGTGTGAAG 240
Qy 181 ACAGCATGGAATCTCTCAAAATGAGCTGTGTAAGTCACTAGTATTTGCAAGGATG 240
Db 241 ACAGCATGGAATCTCTCAAAATGAGCTGTGTAAGTCACTAGTATTTGCAAGGATG 300
Qy 241 TTCAACCGCAATTAAGTATTTGAGAGCGGGAATCAAGCCAGTTATGTTTGTAGTGC 300
Db 301 TTCAACCGCAATTAAGTATTTGAGAGCGGGAATCAAGCCAGTTATGTTTGTAGTGC 360
Qy 301 AAGCTCTGATTTGAGAGCAACAGAGCTTCTTAAAGATCTCAAAAAGAGATGCA 360
Db 361 AAGCTCTGATTTGAGAGCAACAGAGCTTCTTAAAGATCTCAAAAAGAGATGCA 420
Qy 361 ACCAAGATCTGATGAGGAGTAGAGTAGAGATTAAGATGAGTGAAGAAATGAGC 420
Db 421 ACCAAGATCTGATGAGGAGTAGAGTAGAGATTAAGATGAGTGAAGAAATGAGC 480
Qy 421 AAGAGACTGTAAAGTCAACAGGCAACAGCAAGATTTGAAAGGCTTTAAGACTT 480
Db 481 AAGAGACTGTAAAGTCAACAGGCAACAGCAAGATTTGAAAGGCTTTAAGACTT 540
Qy 481 ATGGGGGTTCTGTGTGAGGACACCTTCTGAAGAGAGAGAGAGATGTCAGCCCTTGC 540
Db 541 ATGGGGGTTCTGTGTGAGGACACCTTCTGAAGAGAGAGAGATGTCAGCCCTTGC 600
Qy 541 ATTAAGATTAAGTGTTCCTGCTTGTCTGAGAGATATGAGCTCCCTTACTTTGGGCT 600
Db 601 ATTAAGATTAAGTGTTCCTGCTTGTCTGAGAGATATGAGCTCCCTTACTTTGGGCT 660
Qy 601 CCAGGCTCTCTGCTATTTAATGATCCAAAGTCCAAAGAAATACCTGATGGAATTT 660
Db 661 CCAGGCTCTCTGCTATTTAATGATCCAAAGTCCAAAGAAATACCTGATGGAATTT 720
Qy 661 GATGTGCCAAGTGTGAGAGAGCTTGAATCCATGAGCAAGTATTTGATTTGTC 720
Db 721 GATGTGCCAAGTGTGAGAGAGCTTGAATCCATGAGCAAGTATTTGATTTGTC 780
Qy 721 ATCTGTGTGATGATGATTTGTGATGATGATGATGATGATGATGATGATGATGATG 780
Db 781 ATCTGTGTGATGATGATTTGTGATGATGATGATGATGATGATGATGATGATGATG 840
Qy 781 AACTTAATTCGTACATGAGCTGATGAGAGATCTTGGAGATCTTAATTAAGACAGA 840
Db 841 AACTTAATTCGTACATGAGCTGATGAGAGATCTTGGAGATCTTAATTAAGACAGA 900
Qy 841 TATCAAAATTCCTGAGAGCTGGCTTACCAAGAGCTGAGCGTTGTACAGAGCTTAAT 900
Db 901 TATCAAAATTCCTGAGAGCTGGCTTACCAAGAGCTGAGCGTTGTGTACAGAGCTTAAT 960
Qy 901 GTCAATTTGATATTCCTGAGAGCTGGCTTACCAAGAGCTGAGCGTTGTGTACAGAGCTTAAT 960
Db 961 GTCAATTTGATATTCCTGAGAGCTGGCTTACCAAGAGCTGAGCGTTGTGTACAGAGCTTAAT 1020
Qy 961 TTCTGTGTAAGAAATTAATGTTTCAACGAAGATGGGTGACAAAGGCTTACAGAGATC 1020
Db 1021 TTCTGTGTAAGAAATTAATGTTTCAACGAAGATGGGTGACAAAGGCTTACAGAGATC 1080
Qy 1021 AATCTGTCAGAAATTAATGTTTCAACGAAGATGGGTGACAAAGGCTTACAGAGATC 1080
Db 1081 AATCTGTCAGAAATTAATGTTTCAACGAAGATGGGTGACAAAGGCTTACAGAGATC 1140
Qy 1081 ACCAAGATCTGATGAGGAGTAGAGTAGAGATTAAGATGAGTGAAGAAATGAGC 1140
Db 1141 ACCAAGATCTGATGAGGAGTAGAGTAGAGATTAAGATGAGTGAAGAAATGAGC 1200
Qy 1141 AACAGAAACAAAGGCTGTGGAAGAGAAATTAATCTTGGATGCTTATGATACACTA 1200
Db 1201 AACAGAAACAAAGGCTGTGGAAGAGAAATTAATCTTGGATGCTTATGATACACTA 1260
Qy 1201 CGACTACGAAGAGCGGTGGCTGATCAGCTTGCCTTATGATTAATCTTCCCTGTTTGA 1260
|||||
```



Tue May 27 16:02:20 2003

us-09-805-311-5.std.rge

Page 6

DB	1174	ACGAAACAAAGGAGGACGCCGCGAAGAAGAAATTAAGTTCGATTGTCATTCATCAACA	1233
Db	1201	CGAC	1204
Db	1234	GCAT	1237
RESULT 6			
LOCUS	141	XU68141	1469 bp
DEFINITION		Xenopus laevis XENLB mRNA, complete cds.	linear
VERSION		U68141	
KEYWORDS		U68141.1 GI:1549392	
SOURCE			
ORGANISM		African clawed frog.	
REFERENCE			
AUTHORS		Amphibia: Batrachia: Anura: Mesobatrachia: Pipidae: Pipidae:	
TITLE		Xenopodinae: Xenopus.	
JOURNAL		1 (bases 1 to 1469).	
MEDLINE		Blakova, M., Wu, B., Chi, E., Kim, K.H., Trautman, J.K. and Carroll, D.	
PUBMED		Characterization of FEN-1 from Xenopus laevis. cDNA cloning and	
REFERENCE		J Biol. Chem. 273 (51), 34222-34229 (1998)	
AUTHORS		2 (bases 1 to 1469).	
JOURNAL		Blakova, M., Chi, E., Wu, B., Kim, K.-H. and Carroll, D.	
LOCUS		Xenopus laevis	
DEFINITION		Submitted (2002-AUG-1966) Biochemistry, Univ. of Utah, 50 N. Medical	
VERSION		Drive, Salt Lake City, UT 84132, USA	
KEYWORDS		Locotion/Qualifiers	
SOURCE		1. 1469	
		/organism="Xenopus laevis"	
		/db_xref="taxon:8355"	
		/clone="U68141"	
		/clone_auth="Hemmelt, Brivanlou et al. (1991) Development	
		111, 715-724"	
		/dev_stage="tadpole"	
		/note="also present in oocytes, as determined by	
		sequencing of RT-PCR products"	
		/function="5'-3' exonuclease and a structure-specific	
		endonuclease that removes 5' single-stranded 'flaps';	
		implicated in processing of Okazaki fragments during DNA	
		replication and in base excision repair; 5' nuclease"	
		/note="FEN1 homolog; similar to mammalian Dnae 1 or	
		pombe rad2; similar to Xenopus laevis XENB1a encoded by	
		Genbank Accession Number U64563"	
		/codon_start=1	
		/product="XENB1"	
		/protein_id="FAB08478.1"	
		/db_xref="db:1549392"	
		/db_xref="db:1549392"	
		AAADNDKLNLEAGEEVENETKPKRLIKYVTRKHNEDEKLLPVDVDEADAR	
		ATCAVLAAGVACVAAATADMDATATGCPVALIDTSEAKKLTQEFILNPGD	
		NHEDPACILIGSDYCTETRTGIGPKAIDLIRHKTIEETLINDLKKYIP	
		KEAHLQFDEPVDVADIDTETKEDDEDEGIVATKSGKGFSEDRINCAKLN	
		STQDQADDFGVVDSITTEKSNKYSKTKSKTKSGTGGIPAKTKKQK	
		208 C	
		377 g	
		343 C	
BASE COUNT		461 a	
ORIGIN			
Best Local Match		23.1% Score 326.8; DB 5; Length 1469;	
Matches 615; Conservative		0; Mismatches 437; Indels 6; Gaps 2;	
Db	113	ATGCAATATGGGAATTCAGCGTGGCCACCACTTATCCAGAGTGGCGCGCATGAT	172
Db	90	GAGGACGACGAAGTTGACAGACTACTTCGCGCCCAAAATGACGCGTCGACGACGACGATG	149

Dh	173	CAAGAAGCATGATATAAAAGTAACTTGGTGTGGAAAAGAGCGGTGTGATCTCTATGTG	232
Oy	150	CATATTCAGACTTCGCTATCTGTATGTGGAAAGCAAGCGGTGGAACATCTCAAAATGAGAC	209
Dh	233	CATTTCACCATCTCTCTTCTGCGCTGACAGACAGATGGCA--ATACTGCAACATAGAGCA	289
Oy	210	TGTGTAACTCACTACTGTATTCAGAAAGATGTTTCAACGGCAATTAAGATTACTGTAGAC	264
Dh	290	AGCCCAACCAACACACCTCTCATATGATGGATGATGATGATGATGATGATGATGATGATG	349
Oy	270	GGAATATCAACCCAGTATTATGTTTGGATGGCAACCCCTGCTGATATGAAAGCAAGAGCT	329
Dh	350	TGCACTCAACACCCGTTATATGTGTGTGTAAGGAAACCGGCTTCAATATGAATGTCAGAGTAC	409
Oy	330	TGCTTAAAGATCTCAAAAAGATATGTGTACCAACAAATATGATCTGAGGGCAAGTGAAGT	389
Dh	410	AGCCAAACGATGAGCGTTATGAGCGAGAGCGCAGAAAGTGGTGTGAAGCTCCGAGAGAGC	469
Oy	390	AGGAATTAAGATGTGATGTGAAATATGTGACAGAGAGAGCTGTAAAGTGTCAACAGGCACAA	429
Dh	470	TGCGAAGATATTTGAATCTATGAAATCTTATATGAGAGAGCTGTGATGAGTCAACAGACCA	529
Oy	450	CACGAAATATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	509
Dh	530	CATATGAATGTGAAAGATATCTTACCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG	589
Oy	510	TGAAGCAGACGAAATATGTGACCCCTTGCATTAACGATTAAGTGTTCCTGTGCTTTC	569
Dh	590	TGAAATGATGACCAATGCTGCTGCTTACTTAAATGAAGAAAGATGTATGCTGACGACAC	649
Oy	570	AGAAATGTGTGACTCCCTTACTTGTGGGGCTGACAGCTGCTCTGCTCTCTTAAATGATGC	629
Dh	650	TGAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	689
Oy	630	AATCTTCAAAATATGCTGATATGATGATGATGATGATGATGATGATGATGATGATGATG	709
Dh	710	TGAGCTTAAAGATGCTGATTAATGAGAGGTGCATCTAAACCGGTTGTTTGTAGAAATGCG	769
Oy	690	ACTCAACATGACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	749
Dh	770	CATCATCTACAGACATATGTGGATCTCTGATGATGATGATGATGATGATGATGATGATG	829
Oy	750	CATCAAAAGATATGAGGGGGCAACAGCTGTGAAACTTATGCTCAACATGGCTGCATATGA	809
Dh	830	AATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	869
Oy	810	AATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	889
Dh	890	AGAGATCATGATTAATCTTGAATCTCAAGAAATATCCCATCTCGAGAAATGGCTGTCCAA	949
Oy	870	AGAGATCTGACGCTGTGTCAGAGGCTTAATGTACA--TTGATATCTCTGACATTAAT	926
Dh	950	GGAGGATGAGAGCTTTCTGTGAAACAGAGATGATATATGCTGATATCTATCACTAATA	1009
Oy	927	ATGATCTGACACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	986
Dh	1010	GTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1059
Oy	987	CGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1046
Dh	1070	TGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1129
Oy	1047	AGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1084
Dh	1130	GGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1167

RESULT 7

LOCUS XLU64563

DEFINITION Xenopus laevis 5' nuclelease XENLA mRNA, complete cds.

XLU64563

1461 bp

mRNA

linear

VRF 26-JUL-1999



VERSION U64563.1 GI:1490869  
 KEYWORDS  
 SOURCE African clawed frog.  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 Xenopodinae; Xenopus.  
 REFERENCE  
 1 (bases 1 to 1461)  
 Bibikova, M., Wu, B., Chi, E., Kim, K. H., Trautman, J. K. and Carroll, D.  
 Characterization of FEN-1 from Xenopus laevis. cDNA cloning and  
 role in DNA metabolism  
 J. Biol. Chem. 273 (51), 34222-34229 (1998)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 99069415  
 9852084  
 2 (bases 1 to 1461)  
 Bibikova, M., Chi, E., Wu, B., Kim, K.-H. and Carroll, D.  
 Direct Submission  
 Submitted (19-JUN-1996) Biochemistry U of Utah Med. Sch., 50 N.  
 JOURNAL  
 TITLE  
 Medical Dr., Salt Lake City, UT 84132, USA  
 FEATURES  
 Source  
 1. 1461  
 Location/Qualifiers  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8335"  
 11. 1259  
 /function="Implicated in processing Okazaki fragments  
 during DNA replication and in base excision repair;  
 specifically removes single-stranded 5' sequence"  
 /note="Similar to FEN-1, DnaseIV, MFL, RAD27, RTH1, and  
 rad2"  
 /codon\_start=1  
 /product="5' nuclease xFEN1a"  
 /protein\_id="AAB06176.1"  
 /db\_xref="GI:1490870"  
 /translation="MGHIGIKLINDVAPALIKHDIKSYGRKVVADASMCYIPI  
 AVRODNTLNEGEFTSTSHKMGFRYTRMEHGHKLPYVDDGPPOMKSGELAKRSE  
 RRAEAKLEAEAEBAEVENITEKTRKLYKTKKHNECKKLILKIEHPYDACEE  
 ATCAALVAKAYAAATEDMDALITGTVLVLRHLIASAKLPIQERHLNVLDIGI  
 THEQFDICILGSDYCEITRIGIKRAIDLIRHKTIDELIDIRKAPVPENMIDH  
 KEAKHLFEEVVDIDITELKWIPEDEGLVAFMGKEQFEDIRNGAKYKLVPENMIDH  
 STQGRLDPDFVYKTSKREAEKSAKKKATGCTPGAKFKRK"

BASE COUNT 454 a 293 c 365 g 349 t  
 ORIGIN  
 Query Match 22.7% Score 312.8; DB 5; Length 1461;  
 Best. Local Similarity 53.0%; Pred. No. 2.3e-70;  
 Matches 715; Conservative 0; Mismatches 627; Indels 6; Gaps 2;

Db 468 GTGGAGAACATTGAGAGTTCATTAAGACTTGTCAAGTTACCAAGCAACCAATGAG 527  
 Qy 457 GATTGTAAAGCGCTATTAGACTTATGGGGTCCCTGTTAGAGGACCTTCTGAAGCA 516  
 Db 528 GAGTGGCAAAATTTACTTACCCTTAATGGGCTCTTATGATGATGACCCCTGTAAGCT 587  
 Qy 517 GAAGCAGAAATGGACCCCTTGCATTAAGCATTAAGGTTGGCTGTGTCATCAAGAT 576  
 Db 588 GAGCGACCTGTGTGCTCTTACTTAATAACGAGGAAGATGTCTGCAGCAGCAAGAAC 647  
 Qy 577 ATGGACTCCCTTACTTGTGGGGCTCCACGGTTCCCTGCTATTTAATGATCCAAATTC 636  
 Db 648 ATGATGACCCCTGACCTTGGCACCCCTGCTCTTATGACACCTCAAGCAAGCAAGCT 707  
 Qy 637 AAGAAATATCTGATGAGAAATTTGATGTGCAAGGTTTGGAGGAGCTTGAACCTACC 696  
 Db 708 AAAAAGCTCCATATTCAGAGATTCATTAACCGTGTATTCATTCAGACATAGATACCA 787  
 Qy 697 ATGGACCACTTATGATTTGTGATCTCTGTGTGATGTGATTTGTGATGATCAAA 756  
 Db 768 CATGAACAATTTGTGATCTCTGTATCTTGTGCAAGTGTACTGTGAACCATCCGT 827  
 Qy 757 GGTATGGGGGCGCAACAGCTCTGAACCTTATGTCACATGAGTGCATAGAAAGCATC 816  
 Db 828 GGCATTGGAACCAAAAGAGCCATTGACTGATCGTACGATAGACCATTTGAGAGATC 887  
 Qy 817 TTGGAGATCTTAATTAAGACAGATATCAATTCCTGAGACCTGSCCTTACCAAGAGCT 876  
 Db 888 ATTGATTAATGATGACCTCAAGAAATACCCCTCGGAAATCGCGCAACAGAGGCA 947  
 Qy 877 CGAGGCTTTGTCAGAGAGCTTAA--TGTCACTTGTATTCCTGTGACATTAATGAGCT 933  
 Db 948 AATCATCTTTCTTGTGAACCAAGATAGTGCAGACTGACATCACCGAGCGAATGATC 1007  
 Qy 934 GCACCTGATGAGAGAGGCTTCTAATGTTCTCGTAAAGTAATAGTTTCAACGAAGAT 993  
 Db 1008 GACCCGATGAAGAAGAGATGTTGTCCTTCATGTGTGGGGAAGCAGTATTAGGAAGAT 1067  
 Qy 994 CGGATGACAAAGGCCATAGAGAAGATCAATCTGCCAAGTAATATGCTGCAAGAGCA 1053  
 Db 1068 CGCATACGCAAGCGGGCCAAAGTTGGCCAAAGACCGCAAGGACAGACAGGCGCG 1127  
 Qy 1054 CTCGACTCTTTTTCAGGCACTGCCACATCAGCAGCCGCTAAMAGCAAGCAACT 1113  
 Db 1128 CTCGATGACTTTTTCAACTGACCGGATCCGTTAGCTTAATTAAGAAAGAGGAGAA 1187  
 Qy 1114 TCGATTAACAAGCAAGGACCTGCAAGCAAGAAACAAAGGCTGTGTAAGAAAGAA 1173  
 Db 1188 TCCAAAGATCCGCTTAAGAAAGCAAAACCTGAGGTACCCCTGCAGAAATGTTCAAG 1247  
 Qy 1174 TATCTGTGATGCTGATGTACACTACGACACGAAAGCAGCGGTGCGATGACACTC 1233  
 Db 1248 AGGGATTAATGAATGGGCAATTATTAACTGATGCAATGATTTTATACCGGTTAAG 1307  
 Qy 1234 GCTTATTAATTAATCTCCCTGTTTAACCTGAGACTTGTGTAAGTTGCGCCATGTT 1293  
 Db 1308 GACAAATATGCGGAGTCTTAAATTAAGAAATTTGATTAATTAACCTTATCTTAAG 1367  
 Qy 1294 CAGCTGGGGTAGTATGATGTGTGTAAGAGATGTGGTCCAGATTAACAACTATC 1353  
 Db 1368 TGGGTACGTTTCAATGCTAATGTTGTTTATTGTTAAATTAAGTAAACCAAGTTC 1427  
 Qy 1354 GCTGTTTTTGAAGAAAAAAGAAAAA 1381  
 Db 1428 TGTCTGTGTTGATGAAGAAAAAAGAAAAA 1455  
 RESULT 8  
 AF065397 1357 bp mRNA linear VRT 05-JAN-1999  
 LOCUS  
 DEFINITION Xenopus laevis flap endonuclease 1 (FEN1) mRNA, complete cds.  
 ACCESSION AF065397  
 VERSION AF065397.1 GI:4106357



```
/db_xref="taxon:8355"
/CELL_type="oocyte"
1..1149
/gene="FEN1"
1..1149
/function="FEN1"
/function="structure-specific nuclease"
/codon_start=1
/product="flap endonuclease 1"
/protein_id="BAB8707.1"
/db_xref="GI:2674207.1"
/translation="WGHLGLAKLADVAPAIKREHDIKSYFGRKVALIDSMCIYQELI
AVRODGNLTQNEGEFTSHLMGMYRTIRMEHGIKPYVDEGKPKMGSELAKEE
RAEAEKLELQNEEAGEVENIEKPTKRLVYKQHECKKILTMIGIPYVAPCEAE
ATCAALVAKAGKYAAATEMDALPTGTAVLLRHLTASAKKLPIQEFHLNRVIDIGI
THEOFVLDICILGSDYCEITRIGIPKRAIDLRHKTIDEIDINIDLKYPVENMIL
KEAKHLFEPEVVDITELKWIPEDEGDLVAFMGGEFQSEDIIRNGAKKLKLRNG
STQGRUDDFKVTGSVSSTKREASRSKAKKAKTGTGPKRKRCK"
```

```
BASE COUNT 352 a 247 c 306 g 244 t
ORIGIN
Query Match 22.1%; Score 305.4; DB 5; Length 1149;
Best Local Similarity 55.5%; Pred. No. 1.8e-68;
Matches 632; Conservative 0; Mismatches 501; Indels 6; Gaps 2;
```

```
37 ATGGGCAATCAAGGGTTGAGCAAACTGCGGGGACAAATGCGCCGACGATGACATATAC 156
1 ATGGGAATTCACGTTTGGCCAAACTTATCGAGATGTGGCAGCTGACCTATCAAAAG 60
97 CAGAAATTCGAGACTCTTGGCCGCAAAATGCGCGTGCAGCCGACATGACATATAC 156
61 CATGATATTAAGAGTTCTTGGTTCGAAAGTGCGATGATGCTCCATGTGCAATTTAT 120
157 CAGTTCCGATGTAGTTGGAAGACAGCATGGAACCTCCAAATGACCTGTGTA 216
121 CAGTTCTTATGGGTGAGACAGATGSCA---ACACGCTGCACAAACAGAGTCAA 177
217 GTCACTAGTCTTTGACAGAAATGTTCAACCGACCAATTAAGATTACTGAGCGGAATC 276
178 ACCCAAGCCATCTTATGGGATGATGTTTACCTACCATTCGATGCTAGAGATGGCATC 237
277 AAGCCAGTTTATGTTTGGTGGCAACCTCGATATGAGAAACAGAGCTTCTTAA 336
238 AAACCAAGTATGTGTTGATGGGAACCCGACACATGATGAGCCGCTTAATACCAA 297
337 AGATATCTCAAAAGAGATGATGACCAAAAGATCTGAGGCACTAGAGTAGAGAT 396
298 CGCAGTGAAGCGTAGGCGGAGGACAGAAAGTCTTGAAGCTGCCGAGAGGCTGAGAA 357
397 AAAGATGATTTGAAAATTTAGCAAGAGACTGTAAAGGTGCAAGGACAGCAACGAA 456
358 GTGGAGAAACATTTAGAGAGTTTCACTAAAGACTGTCAAGTTTCAAGAGACATGAG 417
457 GATGTAAAGGCTTATTAAGACTTATGAGGCTTCTGTTGAGAGCAGCCTTCTGAAGA 516
418 GAGTGCAAGAAATTAATTAACCTTAAGGGATCTTATGATGATGACACGCTGAAGT 477
517 GAAGAGAAATGTCAGCCCTTTGGATTAAGATAGGTGCTGCTGCTCAAGAGAT 576
478 GAGGCAAGAGTGTGCTTATGTAAGAGGAGAAAGTGTATGCTCAGACAGAGAGAG 537
577 ATGAGACTCCCTACTTTTGGGGCTCCAGCGTCTCCTCGTCATTTAATGATCCAAATTC 636
538 ATGATGCTGCTGACCTTTGGACCCCTGTCTCTTAAAGGACCTCCACCAAGCAAGCT 597
637 AAGAAATACCTGTGATGAATTTGATGTTGCCAAGGTTTGGAGAGCTTGAATCACC 696
598 AAAAAGCTGCTTATTCAGAGTTCCATTAACCGTCTTCATTCAGACATAGATACACA 657
697 ATGAGCAAGTTCATGATTTGTGATGATGATGAGTGTGAGTGTGATGATGATGATGAT 756
658 CATGAACATTTGTGATCTCTGATCTCTGAGAGTACTACTGTAACCAATCCTCT 717
```

```
QY 757 GGTATCGGGGGGCAACAGACCTCTGAAACTATTCGTAAACATGGGTGCATAGAACATC 816
DB 718 GGCATTGGACCCAAAAGAGCCATTGACCTGATCCGCTAGACATTAAGACATTGACAGATC 777
QY 817 TTGAGAAATCTTAATTAAGACAGATATCAATTCCTAGAGACTGGCTTACCAAGAGCT 876
DB 778 ATTGATTAACATTGACCTCAAGAAATATACCCCTCCGGAATCTGCTGCACAAAGAGGAC 837
QY 877 CGACGCTTCTTCAAGAGCTTAA---TGTCACTTTGATATTCCTGTGCTTAAATGAGCT 933
DB 838 AATCATCTTTTCTTGGACACAGAGTACATGACATGACATACCCGAGAGTGAATC 897
QY 934 GCACCTGATGAGAGGCTCTCATTAAGTTTCTGTTAAAGATTAATGTTTCAACAGAGAT 993
DB 898 GAGCCGAGTGAAGAGAGATGTTGCTTCAATGTGTGGGAGAGAGTAACTTACTGAAAGAT 957
QY 994 CGGCTCAAAAGGCCATGAGAGATCAAAATCTGCAAAAGATTAATTCCTGCAAGAGAA 1053
DB 958 CGCATACGACAGAGGGGCAAAAAGTTGGCCAAAGAACCGGACAGACACAGAGCGCG 1017
QY 1054 CTCGAGTCTCTTTTCAAGCAACTGCGCACACATCAGCAGCCGCTAAAGCGAAGAGACT 1113
DB 1018 CTCGATGACCTTTTCAAGAGTACCCGATCCCTAGCTCACTCACTAAAGAAAGAGAGAA 1077
QY 1114 TCGATTAACAGACAGAGCAGCTGCGAACAAGAAACAAAGCTGCTGGTGAAGAGAA 1172
DB 1078 TCCAAAGATCCGCTTAAGAAAGAACCAAACTGAGAGTACCCTCGCAGAAAGTTCAA 1136
```

RESULT 10  
BC010203 2163 bp mRNA linear ROD 22-OCT-2001  
LOCUS  
DEFINITION Mus musculus, similar to flag structure-specific endonuclease,  
clone MGC:7137 IMAGE:3158176, mRNA, complete cds.  
ACCESSION BC010203  
VERSION BC010203.1 GI:16307327  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (05-JUL-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK  
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [villaloboscm.tmc.edu](mailto:villaloboscm.tmc.edu)  
Villalobos, D.K., Luna, R.A., Hale, S.M., Huliyil, S., Lu, X., Garcia,  
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J.J., Yu, W.,  
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 6 Row: d Column: 19.  
Location/Qualifiers  
1..2163  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="MGC:7137 IMAGE:3158176"  
/tissue\_type="Mammary tumor, C3(1)-Tag model. Infiltrating  
ductal carcinoma. 5 month old virgin mouse."

FEATURES  
source

```

/clone_lib="NCI_GAP_Mam6"
/lab_host="DHI0B"
/notes="Vector: PCMV-SPORT6"
206..1348
/codon_start=1
/product="Similar to flag structure-specific endonuclease"
/protein_id="AAH10203.1"
/db_xref="gi:16307328"
/translation="MGHGLKLLNDVPSAIRNDIKSYFGKRVADIASMSIYQPLI
ATROGQDVLDNBEGETSHLMGFTRTIRIMENETIKVYVFDGKPOLKGETAKSE
RAAEAKOLOQAEAGMEVEKFTKRLVYVKQHNDECKHLISLPYLDPSAE
ASCAALAKAKVYAAATMDMLTFGSPVLRHLTJASBAKLPJOEHLISRVQLEGL
NOEYVDLCILISDYEESIRIGIKRAVDLIQKXSIIEELVRLDPSKPYENMLH
KRAOOLFLEPVLDESEVELKSENEELVYKFMCGKROFSEERIRGVYRLSKSRQG
STOGRLDPEFVYTGSLSSAKRKEPEPKAKKRAKKTGGAGKFRGK"
BASE COUNT      540 a      528 c      619 g      476 t
ORIGIN
Query Match      22.1%; Score 305.4; DB 10; Length 2163;
Best Local Similarity 56.3%; Pred. No. 2e-68;
Matches 642; Conservative 0; Mismatches 481; Indels 18; Gaps 3;
QY 37 ATGGGATCAAGGCTTGACGAACTGCTGGCGGAGCAATGCGCCCAAGCGCATGAGAG 96
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 206 ATGGGAATTCAGCGCTTGCAAACTAATGCTGATGTCGCCCAAGTCCCTGAG 265
QY 97 CAGAACTTCAGAGCTACTTCGCGCCGCAAAATGCGCGTGCAGCGCAGCATGAGCATATATC 156
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 266 AATGACATCAAGAGCTACTTGTGTCGCAAAAGTGGCCATGATGCTCCATGAGCATCTAC 325
QY 157 CAGTTCCTGATGCTTGGAAGAGACAGGACATGSAACCTCACAATGAAGCTGCTGA 216
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 326 CAGTTCCTGATGCTGCTGCTCAG--GGTGGGAGTGTCTCAGACAGAGAGAGGAG 382
QY 217 GTCACATGATTTGCAAGAGATGTTCAACCGACATTAAGTACTGAAACCGGAATC 276
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 383 ACCACCGACCACTGATGATGTCATCTACCGTACCATCCGATGATGAGATGGCATC 442
QY 277 AAGCCATTTATGTTTTGATGAGAGCGCCGCTGATGATGAAGAAACAGAGCTTCTAA 336
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 443 AAGCCTGTGATGCTTTGATGAGCAACACACAGCTAATGTCAGGAGTGGCCAG 502
QY 337 AGATTAACAAAAGAGATGATGCAACCAAGATCTGATGAGGAGAGTAGAGAT 396
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 503 CGGATGAGAGCGCGCGCGCTGAGAAAGCAACTCAGACGAGCTCAGAGGCTGGATG 562
QY 397 AAGATGCGATTAAGAAATGACAGAGAGAGCTGTAAGAGTCAACAAGCAACACAGCA 456
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 563 GAGGAGAGGTGAGAGATTCACACAGAGAGCTGTGAAGGTCAACAAGCAACATGAT 622
QY 457 GATTTAAACGGCTATTAAGACTTATGAGGAGTTCCTGTTGTAAGGACACCTTGTAGACA 516
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 623 GAGTCAAAACCTGCTGATGAGGCTGATCCCTTACCTGATGAGACCCAGCAGAGCA 682
QY 517 GAGCAAAAGTGCAGCGCTTTCATTAAGATTAAGTGTGCTGCTTCTCAAGAT 576
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 683 GAGGCAAGCTGTGCTCGCGCAAGGCTGCAAGTCTATGCTCGCGGCAAGGAGAGAT 742
QY 577 ATGAGACTCCCTTACTTTGGGGCTCCACGCTTCCTTCTGATTTAATGATCAAGTTTC 636
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 743 ATGAGCTGCTTACTTTGGCAGCCCGGCTGATGATGACACTTAATCTCCAGTAGGCC 802
QY 637 AAGAAATTAAGTGTGATGAGATTTGATGTTGCCAAGTTTGGAGAGAGTTAATCTCAC 696
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 803 AAGAGCTGCCCATCCACAGATTCATGAGCGCGCTCTGAGAGAGCTGGCTCTGAGAC 862
QY 697 ATGAGCAGCTTCAATGATTTGTGATCTGTGTGATGTGATGATGATGATGATCAAA 756
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 863 CAGAGAGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 922
QY 757 GGTATCGGGGGGCAAGAGCTGTAAGACTTATTCATGATGATGATGATGATGATGATGATGAT 816
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 923 GGCATTGGGCGCAAGCGGCTGTGATCTCATCAGAAACATTAAGAGATGAGAGATGATC 982

```

```

QY 817 TTGGACATCTTATTAAGACAGATATCAATTCCTGAGAGCTGGCTTACCAAGAGAT 876
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 983 GTGAGCGGCTGAGACCCAGCAGATACCCGCTGAGAGCTGCTCCACAGAGAGCC 1042
QY 877 CCACGCTTTGTCAGAGAGCTTATG---TCACATTTGATTTCTGAGCTAAATGAGCT 933
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1043 CAGAGAGCTTCTCTGAGAGCAGAACTACTGAGCCOAGAGCTGTGAGAGCTGAATGAGAG 1102
QY 934 GCACCTGATGAGAGAGGCTCTCATAGTTTCTGTTGTTAAAGATTAATGTTTCAACGAGAT 993
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1103 GAACCAAAAGAGAGAGTTGGTCAATTTATGTTGTGTAAGAGAGTTTCTGAGAGAG 1162
QY 994 CGGCTGACCAAGGCGCATAGAGAGATCAATCTGCCAAGATTAATGCTGCAGAGAGAGA 1053
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1163 CGAATTCGAGAGTGGGGTCAAGCGGCTGAGTAAGAGCCGCCAGGCGACAGCCAGAGAGCG 1222
QY 1054 CTCGAGTCTTTTTCAGCCACCTGCTCCACACATGACAGCGCTTAAGAGAGAGACT 1113
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1223 CTCGAGATTTCTTCAGGCTGACAGCTCCTACTCTCCTAGC-----TAAGCGC 1270
QY 1114 TCGATTAACAAGCAGAGCAGCTCGCAACAAGAAACAAAGCGCTGTGCAAGAGAGAA 1173
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1271 AAGGAGCCAGAACCCAGAGGCGCTGCTAAGAAAGAGAAAGCAAGCTGCGGAGCGGAG 1330
QY 1174 T 1174
    |
Db 1331 T 1331

```

RESULT 11  
AC026761  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
SUBMISSION  
On Aug 21, 2001 this sequence version replaced g1:10937955.  
-----Genome Center  
Center: Harvard Partners Genome Center  
Center Code: HPGC  
Web site:  
http://wchanming.bwh.harvard.edu:9088/hpcpg/j5p/hpcpg/sequence/mous  
e.html  
Contact: gntm@aceped.bwh.harvard.edu  
-----Summary Statistics  
Center project name: ACD  
Sequencing vector: pUC18; L08752  
Chemistry: Dye-terminator Big Dye; 100%  
\*Consensus quality: 156712 at least Q20  
\*Consensus quality: 152694 at least Q30  
\*Consensus quality: 147247 at least Q40  
Estimated insert size: agarose-FP - N/A  
\*Estimated insert size: 165608 - sum-of-contigs

Quality coverage:	agarose-FP - N/A
Quality coverage:	4 x in Q20 bases; sum-of-contigs estimation

NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	34556: contig of 34556 bp in length
34557	34576: gap of unknown length
34577	62727: contig of 28151 bp in length
62728	62747: gap of unknown length
62748	81216: contig of 18465 bp in length
81217	81236: gap of unknown length
81237	92870: contig of 11634 bp in length
92871	92890: gap of unknown length
92891	103627: contig of 10737 bp in length
103628	1036447: gap of unknown length
103648	1146597: contig of 11050 bp in length
114698	114717: gap of unknown length
114718	123385: contig of 8668 bp in length
123386	123405: gap of unknown length
123406	131015: contig of 8410 bp in length
131016	131835: gap of unknown length
131816	132037: contig of 202 bp in length
131836	132057: gap of unknown length
132038	138473: contig of 6416 bp in length
132058	138473: contig of 6416 bp in length
138474	138493: gap of unknown length
138494	144023: contig of 5530 bp in length
144024	144023: gap of unknown length
144025	144824: contig of 881 bp in length
144825	144844: gap of unknown length
144925	145191: contig of 247 bp in length
144926	145211: gap of unknown length
145192	145211: gap of unknown length
145212	148414: contig of 3203 bp in length
148415	148434: gap of unknown length
148435	151539: contig of 3105 bp in length
151539	151559: gap of unknown length
151560	154468: contig of 2809 bp in length
154469	154488: gap of unknown length
154489	156193: contig of 1705 bp in length
156194	156213: gap of unknown length
156214	157744: contig of 1531 bp in length
157745	157764: gap of unknown length
157765	159461: contig of 1697 bp in length
159462	159481: gap of unknown length
159482	160346: contig of 865 bp in length
160347	160366: gap of unknown length
160367	161824: contig of 1458 bp in length
160375	161844: gap of unknown length
161845	163683: contig of 1819 bp in length
163684	163683: gap of unknown length
163684	164031: contig of 348 bp in length
164032	164051: gap of unknown length
164052	164700: contig of 649 bp in length
164701	164720: gap of unknown length
164721	165088: contig of 1368 bp in length

## FEATURES

### Location/Qualifiers

```
misc_feature /note="assembly_name:Contig79" 81237. 92870
misc_feature /note="assembly_name:Contig78" 92891. 103627
misc_feature /note="assembly_name:Contig77" 103648. 114657
misc_feature /note="assembly_name:Contig76" 114718. 123385
misc_feature /note="assembly_name:Contig75" 123406. 131815
misc_feature /note="assembly_name:Contig74" 131836. 132057
misc_feature /note="assembly_name:Contig73" 132058. 138473
misc_feature /note="assembly_name:Contig72" 138494. 144033
misc_feature /note="assembly_name:Contig71" 144044. 144934
misc_feature /note="assembly_name:Contig70" 144945. 145191
misc_feature /note="assembly_name:Contig69" 145212. 148414
misc_feature /note="assembly_name:Contig68" 148435. 151539
misc_feature /note="assembly_name:Contig67" 151560. 154468
misc_feature /note="assembly_name:Contig66" 154489. 156193
misc_feature /note="assembly_name:Contig65" 156214. 157744
misc_feature /note="assembly_name:Contig64" 157765. 159461
misc_feature /note="assembly_name:Contig63" 159482. 160346
misc_feature /note="assembly_name:Contig62" 160367. 161824
misc_feature /note="assembly_name:Contig61" 161845. 163663
misc_feature /note="assembly_name:Contig60" 163684. 164031
misc_feature /note="assembly_name:Contig59" 164052. 164700
misc_feature /note="assembly_name:Contig58" 164721. 166068
misc_feature /note="assembly_name:Contig57" 166089. 167447
misc_feature /note="assembly_name:Contig56" 167468. 168826
misc_feature /note="assembly_name:Contig55" 168847. 170205
misc_feature /note="assembly_name:Contig54" 170226. 171584
misc_feature /note="assembly_name:Contig53" 171603. 172961
misc_feature /note="assembly_name:Contig52" 172980. 174338
misc_feature /note="assembly_name:Contig51" 174357. 175715
misc_feature /note="assembly_name:Contig50" 175734. 177092
misc_feature /note="assembly_name:Contig49" 177111. 178469
misc_feature /note="assembly_name:Contig48" 178488. 179846
misc_feature /note="assembly_name:Contig47" 179865. 181223
misc_feature /note="assembly_name:Contig46" 181242. 182600
misc_feature /note="assembly_name:Contig45" 182617. 183975
misc_feature /note="assembly_name:Contig44" 183992. 185350
misc_feature /note="assembly_name:Contig43" 185367. 186725
misc_feature /note="assembly_name:Contig42" 186742. 188100
misc_feature /note="assembly_name:Contig41" 188115. 189480
misc_feature /note="assembly_name:Contig40" 189495. 190860
misc_feature /note="assembly_name:Contig39" 190875. 192250
misc_feature /note="assembly_name:Contig38" 192265. 193640
misc_feature /note="assembly_name:Contig37" 193655. 195030
misc_feature /note="assembly_name:Contig36" 195040. 196415
misc_feature /note="assembly_name:Contig35" 196425. 197800
misc_feature /note="assembly_name:Contig34" 197810. 199185
misc_feature /note="assembly_name:Contig33" 199195. 200570
misc_feature /note="assembly_name:Contig32" 200580. 201955
misc_feature /note="assembly_name:Contig31" 201965. 203340
misc_feature /note="assembly_name:Contig30" 203350. 204725
misc_feature /note="assembly_name:Contig29" 204735. 206110
misc_feature /note="assembly_name:Contig28" 206120. 207495
misc_feature /note="assembly_name:Contig27" 207500. 208875
misc_feature /note="assembly_name:Contig26" 208885. 210260
misc_feature /note="assembly_name:Contig25" 210270. 211645
misc_feature /note="assembly_name:Contig24" 211655. 213030
misc_feature /note="assembly_name:Contig23" 213040. 214415
misc_feature /note="assembly_name:Contig22" 214425. 215800
misc_feature /note="assembly_name:Contig21" 215810. 217185
misc_feature /note="assembly_name:Contig20" 217195. 218570
misc_feature /note="assembly_name:Contig19" 218580. 219955
misc_feature /note="assembly_name:Contig18" 219965. 221340
misc_feature /note="assembly_name:Contig17" 221350. 222725
misc_feature /note="assembly_name:Contig16" 222735. 224110
misc_feature /note="assembly_name:Contig15" 224120. 225495
misc_feature /note="assembly_name:Contig14" 225500. 226875
misc_feature /note="assembly_name:Contig13" 226885. 228260
misc_feature /note="assembly_name:Contig12" 228270. 229645
misc_feature /note="assembly_name:Contig11" 229655. 231030
misc_feature /note="assembly_name:Contig10" 231040. 232415
misc_feature /note="assembly_name:Contig9" 232425. 233800
misc_feature /note="assembly_name:Contig8" 233810. 235185
misc_feature /note="assembly_name:Contig7" 235195. 236570
misc_feature /note="assembly_name:Contig6" 236580. 237955
misc_feature /note="assembly_name:Contig5" 237965. 239340
misc_feature /note="assembly_name:Contig4" 239350. 240725
misc_feature /note="assembly_name:Contig3" 240735. 242110
misc_feature /note="assembly_name:Contig2" 242120. 243495
misc_feature /note="assembly_name:Contig1" 243500. 244875
misc_feature /note="assembly_name:Contig0" 244885. 246260
misc_feature /note="assembly_name:Contig0" 246270. 247645
misc_feature /note="assembly_name:Contig0" 247655. 249030
misc_feature /note="assembly_name:Contig0" 249040. 250415
misc_feature /note="assembly_name:Contig0" 250425. 251800
misc_feature /note="assembly_name:Contig0" 251810. 253185
misc_feature /note="assembly_name:Contig0" 253195. 254570
misc_feature /note="assembly_name:Contig0" 254580. 255955
misc_feature /note="assembly_name:Contig0" 255965. 257340
misc_feature /note="assembly_name:Contig0" 257350. 258725
misc_feature /note="assembly_name:Contig0" 258735. 260110
misc_feature /note="assembly_name:Contig0" 260120. 261495
misc_feature /note="assembly_name:Contig0" 261500. 262875
misc_feature /note="assembly_name:Contig0" 262885. 264260
misc_feature /note="assembly_name:Contig0" 264270. 265645
misc_feature /note="assembly_name:Contig0" 265655. 267030
misc_feature /note="assembly_name:Contig0" 267040. 268415
misc_feature /note="assembly_name:Contig0" 268425. 269800
misc_feature /note="assembly_name:Contig0" 269810. 271185
misc_feature /note="assembly_name:Contig0" 271195. 272570
misc_feature /note="assembly_name:Contig0" 272580. 273955
misc_feature /note="assembly_name:Contig0" 273965. 275340
misc_feature /note="assembly_name:Contig0" 275350. 276725
misc_feature /note="assembly_name:Contig0" 276735. 278110
misc_feature /note="assembly_name:Contig0" 278120. 279495
misc_feature /note="assembly_name:Contig0" 279500. 280875
misc_feature /note="assembly_name:Contig0" 280885. 282260
misc_feature /note="assembly_name:Contig0" 282270. 283645
misc_feature /note="assembly_name:Contig0" 283655. 285030
misc_feature /note="assembly_name:Contig0" 285040. 286415
misc_feature /note="assembly_name:Contig0" 286425. 287800
misc_feature /note="assembly_name:Contig0" 287810. 289185
misc_feature /note="assembly_name:Contig0" 289195. 290570
misc_feature /note="assembly_name:Contig0" 290580. 291955
misc_feature /note="assembly_name:Contig0" 291965. 293340
misc_feature /note="assembly_name:Contig0" 293350. 294725
misc_feature /note="assembly_name:Contig0" 294735. 296110
misc_feature /note="assembly_name:Contig0" 296120. 297495
misc_feature /note="assembly_name:Contig0" 297500. 298875
misc_feature /note="assembly_name:Contig0" 298885. 300260
misc_feature /note="assembly_name:Contig0" 300270. 301645
misc_feature /note="assembly_name:Contig0" 301655. 303030
misc_feature /note="assembly_name:Contig0" 303040. 304415
misc_feature /note="assembly_name:Contig0" 304425. 305800
misc_feature /note="assembly_name:Contig0" 305810. 307185
misc_feature /note="assembly_name:Contig0" 307195. 308570
misc_feature /note="assembly_name:Contig0" 308580. 309955
misc_feature /note="assembly_name:Contig0" 309965. 311340
misc_feature /note="assembly_name:Contig0" 311350. 312725
misc_feature /note="assembly_name:Contig0" 312735. 314110
misc_feature /note="assembly_name:Contig0" 314120. 315495
misc_feature /note="assembly_name:Contig0" 315500. 316875
misc_feature /note="assembly_name:Contig0" 316885. 318260
misc_feature /note="assembly_name:Contig0" 318270. 319645
misc_feature /note="assembly_name:Contig0" 319655. 321030
misc_feature /note="assembly_name:Contig0" 321040. 322415
misc_feature /note="assembly_name:Contig0" 322425. 323800
misc_feature /note="assembly_name:Contig0" 323810. 325185
misc_feature /note="assembly_name:Contig0" 325195. 326570
misc_feature /note="assembly_name:Contig0" 326580. 327955
misc_feature /note="assembly_name:Contig0" 327965. 329340
misc_feature /note="assembly_name:Contig0" 329350. 330725
misc_feature /note="assembly_name:Contig0" 330735. 332110
misc_feature /note="assembly_name:Contig0" 332120. 333495
misc_feature /note="assembly_name:Contig0" 333500. 33
```

```

Db 11711 CGCAGTGGAGAGCGCGCGGAGGCTGAGAAAGCACTCCAGAGGCTCAGAGGCTGGGATG 11770
Oy 397 AAGATGCGATTGAAAAATTGAGCAGAGAGACTGTAAGGCTACACAAAGCAACAGCAA 456
Db 11771 GAGGAGGAGGTGGAGAAAGTTCCACCAAGAGGCTGCTGACAGTCCACCAAGCAACAGTAT 11830
Oy 457 GATTGTAAAGCGCTATTAAAGCTATTAGGGGCTTCTGTTGTAAGAGCACTTCTGAAAGCA 516
Db 11831 GAGTGTAAAGCACTGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 11890
Oy 517 GAGGAGAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 576
Db 11891 GAGGAGAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 11950
Oy 577 ATGAGCTGCTTACTTTTGGGGGCTCCAGGCTTCTTCTGATCTTAAATGATCCAGTTCC 636
Db 11951 ATGAGCTGCTTACTTTTGGGGGCTCCAGGCTTCTTCTGATCTTAAATGATCCAGTTCC 12010
Oy 637 AAGAAATACCTCTGATGAGATTTGATGTTGCAAGGCTTTGAGAGGCTTGAACCTCACC 696
Db 12011 AAGAAATACCTCTGATGAGATTTGATGTTGCAAGGCTTTGAGAGGCTTGAACCTCACC 12070
Oy 697 ATGAGGCTGCTTACTTTTGGGGGCTCCAGGCTTCTTCTGATCTTAAATGATCCAGTTCC 756
Db 12071 CAGGAGGAGTTTGGAGTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 12130
Oy 757 GGTATCGGGGGGCAAGAGCTCTGAAACTTATTCGTCAACATGAGTGCATAGAAAGCATC 816
Db 12131 GGTATCGGGGGGCAAGAGCTCTGAAACTTATTCGTCAACATGAGTGCATAGAAAGCATC 12190
Oy 817 TTGAGATCTTAAATGAGAGAGATGATCAATTCCTGAGAGGCTGGCTTACCAAGAGCT 876
Db 12191 TTGAGATCTTAAATGAGAGAGATGATCAATTCCTGAGAGGCTGGCTTACCAAGAGCT 12250
Oy 877 CGAGCTGTTGTTCAAGAGAGCTTAATG--TCATCTTGAATTTCTGAGCTTAAATGACT 933
Db 12251 CGAGCTGTTGTTCAAGAGAGCTTAATG--TCATCTTGAATTTCTGAGCTTAAATGACT 12310
Oy 934 GCACCTGAGAGAGGCTCTCATTAATTTCTGTTAAAGATTAATGTTTCAAGAGAT 993
Db 12311 GCACCTGAGAGAGGCTCTCATTAATTTCTGTTAAAGATTAATGTTTCAAGAGAT 12370
Oy 994 CGGAGTGAAGAGGCTATGAGAGATCAATCTGCCAACAATTAATCTGTCAGAGAGAA 1053
Db 12371 CGAATTCGAGTGGGCTCAAGAGGCTGAGTGAAGAGGCTGAGGCTGAGGCTGAGGCT 12430
Oy 1054 CTGAGTCTCTTTTCAAGAGGCTGAGTGAAGAGGCTGAGTGAAGAGGCTGAGTGAAGAG 1113
Db 12431 CTGAGTCTCTTTTCAAGAGGCTGAGTGAAGAGGCTGAGTGAAGAGGCTGAGTGAAGAG 12478
Oy 1114 TCGGATTAAGAGAGGCTGAGTGAAGAGGCTGAGTGAAGAGGCTGAGTGAAGAGAGAA 1173
Db 12479 AAGAGGAGAGAGGCTGAGTGAAGAGGCTGAGTGAAGAGGCTGAGTGAAGAGAGAG 12538
Oy 1174 T 1174
Db 12539 T 12539

```

## RESULT 12

```

LOCUS      HSRAD2      1757 bp      mRNA      linear      PRI 07-DEC-1998
DEFINITION Homo sapiens mRNA for flap endonuclease-1.
ACCESSION  X76771
VERSION    X76771.1 GI:1905802
KEYWORDS   feat gene; flap endonuclease-1.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1757)
AUTHORS   Watts,F.

```

```

TITLE      Direct Submission
JOURNAL    Submitted (16-DEC-1993) F. Watts, University of Sussex, School of
Biological Sciences, Falmer, Brighton BN1 9QG, UK
REFERENCE  2 (bases 1 to 1757)
AUTHORS   Murray,J.M., Navesoli,M., al-Hatithy,R., Sheldrick,K.S.,
            Lehmann,A.R., Carr,A.M. and Watts,F.Z.
TITLE      Structural and functional conservation of the human homolog of the
            Schizosaccharomyces pombe rad2 gene, which is required for
            chromosome segregation and recovery from DNA damage
JOURNAL    Mol. Cell. Biol. 14 (7), 4878-4888 (1994)
MEDLINE    94277093
FEATURES   location/Qualifiers
            source      1..1757
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
            gene        368..1510
                        /gene="fen1"
                        /gene="fen1"
                        /gene="fen1"
            CDS          368..1510
                        /product="flap endonuclease-1"
                        /protein_id="CAA54166.1"
                        /db_xref="GI:3980293"
                        /translation="MGIOGLAKLADVAPSPNIRENDIKSYFGRKVAIDMSITYPLI
                        AVRGSDVLQNEGETSHLMGFRTIRRMENGIKPYVEDKPKQLSGLKARSE
                        RRAEARKIQDQQAAGAEDEVEKFTKRLVYVTOHNDCKHLISLNGIPYLDAPSEAE
                        ASCNALVRAKGRVYAAATDEMDCLTEGSPVIMRLHLSSEAKKLPIDFHLSTRLOELG
                        NOEQFVLDLILGSDYCEISIGIPKRAVDLIDKHSIEIVRLDPNRYPYPENMLH
                        KEAHOLFLEPEVLDESEVLEKMSPEDEILIKFCKGSEKSEIRISGRISRLSKSROG
                        STQGRIDDFEYVYGLSSAKRKEPEPGRSTKRKAATGAACKFPRK"
BASE COUNT  418 a      443 c      542 g      354 t
ORIGIN
Query Match      22.0%; Score 304.2; DB 9; Length 1757;
Best Local Similarity 54.6%; Pred. No. 4e-68;
Matches 675; Conservative 0; Mismatches 553; Indels 9; Gaps 3;
Oy 2 GACCAGCGCTGCGGCGACAGCGCGGAGAGAGATGAGGCTGAGGCTGAGGCTGAGGCT 61
Db 333 GTCCCAAGGCGCATCTATCCCTCTCTGTGTTGCCATGGAATTCAGAGGCTGGCCAAAC 392
Oy 62 TGCCTGGGCGGCAATGCGGCCCAAGCGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 121
Db 333 TAATTCGATGTTGCCCCCGAGTCCATCCGGAGATGACATCAAGAGCTACTTTGGCC 452
Oy 122 GCAAAATCGCCGTCGAGCGCCAGCATGAGCATATACAGTTCTGATTTGTTGAGAGGA 181
Db 453 GTAAGGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 511
Oy 182 CAGGATGSAAGACCTGCAAAATGAGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAG 241
Db 512 --GGTGGGATGCTGTCAGAGATGAGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAG 569
Oy 242 TCAACCGGCAATTAAGATTACTGGAAGCGGAGTCAAGCCAGTTTATGTTTGTGATGGA 301
Db 570 TCTACCGGCAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 629
Oy 302 AGCCTCGTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
Db 630 AGCGGCCACAGCTCAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 689
Oy 362 CCAAGATCTGACTGAGGAGCTGAGAGTGAAGATGAAGATGAGATGAGATGAGATGAGATGAG 421
Db 690 AAGAGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 749
Oy 422 AGAGAGCTGTAAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
Db 750 AGGGCTGAGTGAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 809
Oy 482 TGGGGGTTCTGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
Db 810 TGGGATTCCTTTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 869

```

```

QY 542 TAAACATTAAGGTGCTGCTGCTTCAGAAAGATATGAGTCCCTTACTTTTG6G6CTC 601
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 870 AGGCTTG6CAAGCTGATGCTGCGTACCGAGGACATGAGTGCCTCAGCTCGGCGCC 929
QY 602 CAGGTTCTCTGCTGCTATTAATGATTCACAGTTCAGAAATATCTGTGATGATTTG 661
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 930 CTGTGCTAATGGAGACCTGACTGTGCGAGTGAAGCAAAAAGCTGCAATCCAGSANTTC 989
QY 662 ATGTGCGCAAGTTTGGAGAGAGCTTGAACCTCACCATGAGACCATCTATTTATTTGCA 721
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 990 ACCTGAGCGGATTTGCGAGAGCTGGGCTTAACACAGAAAGTTGTGTGATGATGCA 1049
QY 722 TCTGTGTGATGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 781
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1050 TCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1109
QY 782 AACTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 841
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1110 ACCTGATCCGAGAGCAGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAG 1169
QY 842 ATCAAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1170 ACCCTGTGCGAGAAATTTGCTTCACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1229
QY 902 ---TCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1230 TGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1289
QY 959 GTTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1018
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1290 AGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1349
QY 1019 TCAAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1078
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1350 TGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1409
QY 1079 CCAC---CACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1135
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1410 GCTGATCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1469
QY 1136 CTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1195
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1470 CAAGAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1529
QY 1196 AACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1232
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1530 CTCTTCACCCAGAGATATTTGCTTGTGACCTT 1566

RESULT 13
BC000323      2031 bp      mRNA      linear      PRI 12-JUL-2001
LOCUS        Homo sapiens, flap structure-specific endonuclease 1, clone
DEFINITION   MGC:8478 IMAGE:2821792, mRNA, complete cds.
ACCESSION   BC000323
VERSION     BC000323.1 GI:12653112
KEYWORDS    MGC.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2031)
AUTHORS     Strausberg, R.
TITLE        Direct Submission
SUBMITTER   Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
REMARK      NIH-MGC project URL: http://mgc.ncl.nih.gov
CONTACT     MGC help desk
            Email: cgabs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory

```

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@ncl.nih.gov
Shechenko, Y., Welberry, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-T., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Mastello, C., Mastriani, S.D., McCloskey, J.C.,
McGowan, J., Pearson, R., Snyder, B., Stancil, P., Thomas, P.J.,
Tongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L., H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 1 Row: 9 Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 1905802.

FEATURES
    source
        1..2031
        /organism="Homo sapiens"
        /db_xref="LocusID:2237"
        /clone="MGC:8478 IMAGE:2821792"
        /tissue_type="lung, small cell carcinoma"
        /clone_lib="NIH-MGC-7"
        /lab_host="DH10B-R"
        /note="Vector: pOTB"
        139..1281
        /product="flap structure-specific endonuclease 1"
        /protein_id="AAH00323.1"
        /db_xref="GI:12653113"
        /translation="MGIOLAKLIDAVPSAIRBNDIKSYREKRAVIDASISYIOPILI
        AYRGSGVLDNEGGTTSHLMGFFRTIMNGKIPYVDPGKPKQIKSEILKRSF
        BRABERQLODQAGABQDEVEKTRKLVKTRKONHDEKHLISMGPIYDASBAP
        ASCAVALKAKVYAAATVEDMDCTFESPLVTHLHLSAEKRLPDIOPHLSLDELGL
        NQEOEVLICLLSDYCESLRIGIRAVADLQKRSLEIVRLDPKRYVPVNWHL
        KEAHLFLPEVDELSPESVLEKMSPEBEELIFEMGEGKQFSEERTRSSVKRLSKSROG
        STQORLDLDFKVTGSLSSAKRKEPEKSGTSKKRKYTAAGKFKRKG"

BASE COUNT      533 a      465 c      589 g      444 t
ORIGIN
Query Match      22.0%; Score 304.2; DP 9; Length 2031;
Best Local Similarity 54.6%; Pred. No. 4.1e-68;
Matches 675; Conservative 0; Mismatches 553; Indels 9; Gaps 3;
QY 2 GACCCAGCGTCCGCGCAGCCGCCGAGACGAGATGGCATCAAGAGTTTGACGAAC 61
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 104 GTCCCAAGGCGAGTCATCTCCCTGCTGTGTCGATGGGAATTCAGAGCGTGGCAAC 163
QY 62 TGTGTGCGGACATATCCGCCCAAGCGATGAAGAGCAGAACTTCAGAGAGTTCGGCC 121
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 TAATGCTGATGTGCGCCCGCAGTCCATCCGGAATGACATCAAGATCTTGTGCGC 223
QY 122 GCAAAATGCGCGTCCAGCCGACGATGAGCATATACAGTCTGTGATGTGTGGAAGA 181
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 GTAAGGTGCGCATGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 282
QY 182 CAGCGATGGAATCTCCAAATGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 241
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 -GGTGGGATGTGCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 340
QY 242 TCAACCGGCAATTAAGATTACTGGAAGCGGAGATCAAGCAGTTATGTTTGAATGGA 301
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 TCTACCGCACCACTTGTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 400
QY 302 AGCTTCCTGATTAAGAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 401 AGCGGCGACAGCTCAAGTCAGGCGAGCTGGCCAAACGAGTGAAGCGGCGGCTGAGAGCAG 460
QY 362 CCAAAGATCTGAGTGAAGGAGTGAAGGATTAAGATGCGATTGAAGAAATTTAGAGA 421

```



```

Db 461 AGAAGCAGCTGCAGCAGGCTCAGCTGCTGGGCCGACGAGGAGGTGGAAAAATTCACCTA 520
Oy 422 AGAGACTGTAAAGTCAACAGCAACACAGAGATTTGTAACGGCTATTAAAGACTTA 481
Db 521 AGCGCTGTGTGAAGTCTACAGCAGCAATGATGAGCAACATCTGCTGACGCTCA 580
Oy 482 TGGGGTTCCTGTTGTAGAGCAGCTCTTGAACACAGAGATTTGAGCCCTTTGA 541
Db 581 TGGGCACTCCCTTATCTGTATGACACCCAGTAGAGCAGAGCCAGCTGTGCTGCTGTA 640
Oy 542 TAAACGATTAAGTGTCTGCTGCTTCAAGATATGACATCCCTTACTTTGGGGCTC 601
Db 641 AGCGTGCACAAAGTCTATGCTGCGGTACGAGGACATGATGATCCCTCACCTTGCGACGC 700
Oy 602 CAGGTTCCCTGCTGATTTAATGATCCAAAGTTCCAAAGAAATACCTGATGAGAAATTG 661
Db 701 CTGTGCTAATGCGACACCTCAGCTGCGAGGACCAAAAAAGCTGCTCAATCCAGAAATTC 760
Oy 662 ATGTTGCCAAGGTTTGGAGGAGCTTGAACCTCAACATGACAGCAATGATTTGTGCA 721
Db 761 ACCTGACCGGATTTCTCAGAGAGCTGGGCTGAACAGAGAGATTTGTGATCTGTGCA 820
Oy 722 TCCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
Db 821 TCCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 880
Oy 782 AACTTATTCGTCAACATGAGTGTCAAGAAAGCATCTTGAAGATCTTAAAGACAGAT 841
Db 881 ACCATCATCCAGAAACACAAAGATGAGATGATGATGATGATGATGATGATGATGATGAT 940
Oy 842 ATCAATTCCTGAGAGCTGCGCTTACCAAGATGATGATGATGATGATGATGATGATGATG 901
Db 941 ACCCTGTGCGAAGAAATTTGGCTCCACAGAGAGCTCACCAGCTCTTCTTGAGACCTGACG 1000
Oy 902 ---TCACATGATGATTTCTGAGCTAAATGATGATGATGATGATGATGATGATGATGAT 958
Db 1001 TGCTGAGCCAGAGATCTGTGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1060
Oy 959 GTTTCCTGTGTAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018
Db 1061 AGTTCATGTGTGTGTAAGAAAGATTTCTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1120
Oy 1019 TCAATATCTGCCAGATTAATTCGTGCAAGAGAGCTCGAGTCTTTTCAAGCCACTG 1078
Db 1121 TGAGTAAAGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1180
Oy 1079 CCAC---CACATCAGACCGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1135
Db 1181 GCTCAGCTCTTCTGCTAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1240
Oy 1136 CTGCGAAGCAAGAAACAAAGGCTGTGTGAAGAAAGAAATATCTTGGATGCTTATGATAC 1195
Db 1241 CAAAGACTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1300
Oy 1196 AACTAGCACTAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1232
Db 1301 CTCCTTCAACCCAGCAATATTTGCCGCTTGTACCTT 1337

```

```

RESULT 14
AP000591 138627 bp DNA linear HTG 30-MAY-2000
LOCUS Homo sapiens chromosome 11 clone CMB9-22K5 map 11q13, WORKING DRAFT
DEFINITION
AP000591
ACCESSION AP000591.3 GI:8118797
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT
KEYWORDS Homo sapiens DNA, Clone::CMB9-22K5.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 138627)

```

AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,  
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,  
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,  
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
 Submitted (12-Oct-1999) Masahira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,  
 Japan (E-mail: hattori@gs.c.riken.go.jp,  
 URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923,  
 Fax: 81-42-778-9924)  
 On May 31, 2000 this sequence version replaced gi:6997481.  
 ----- Genome Center  
 Center: RIKEN Genomic Sciences Center (GSC)  
 Center code: RIKEN  
 Web site: http://hgp.gsc.riken.go.jp/  
 Contact: hattori@gs.c.riken.go.jp  
 ----- Project Information  
 Center project name: HumDrafl1  
 Center clone name: CMB9-22K5  
 ----- Summary Statistics  
 Sequencing vector: PCR products: 100% of reads  
 Chemistry: Dye-terminator ET-amersham: 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 122298 bases at least Q40  
 Consensus quality: 130648 bases at least Q30  
 Consensus quality: 134094 bases at least Q20  
 Insert size: 135927; sum-of-ctrls  
 Quality coverage: 4.01x in Q20 bases; sum-of-ctrls  
 -----  
 NOTE: This is a 'working draft' sequence. It currently consists of  
 28 contigs. The true order of the pieces is not known and their  
 order in this sequence record is arbitrary. Gaps between the  
 contigs are represented as runs 'N', but the exact sizes of the gaps  
 are unknown. This record will be updated with the finished sequence  
 as soon as it is available and the accession number will be  
 preserved  
 1  
 20827 20726 contig of 13840 bp in length  
 34767 46065 contig of 11298 bp in length  
 46166 53516 contig of 7351 bp in length  
 53617 61065 contig of 7449 bp in length  
 61166 67998 contig of 6833 bp in length  
 68099 74526 contig of 6428 bp in length  
 74627 81299 contig of 6673 bp in length  
 81400 87316 contig of 5917 bp in length  
 87417 91541 contig of 4125 bp in length  
 91642 95731 contig of 4090 bp in length  
 95832 101128 contig of 5257 bp in length  
 101229 104948 contig of 3720 bp in length  
 105049 110306 contig of 5268 bp in length  
 110407 112969 contig of 2583 bp in length  
 113070 116484 contig of 3415 bp in length  
 115885 119599 contig of 3015 bp in length  
 119700 121881 contig of 2182 bp in length  
 121982 122703 contig of 722 bp in length  
 122804 125566 contig of 2763 bp in length  
 125667 126845 contig of 1179 bp in length  
 126946 129349 contig of 2404 bp in length  
 129450 131268 contig of 1819 bp in length  
 131369 132565 contig of 1197 bp in length  
 132666 134031 contig of 1366 bp in length  
 134132 136176 contig of 2045 bp in length  
 136277 137418 contig of 1132 bp in length  
 137519 138627 contig of 1109 bp in length  
 Sequence updated (20-Oct-1999).  
 Sequence updated (26-May-2000).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 28 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is



- \* arbitrary gaps between the contigs are represented as
- \* runs of N, but the exact sizes of the gaps are unknown.
- \* This record will be updated with the finished sequence
- \* as soon as it is available and the accession number will
- \* be preserved.

[illegible]

## FEATURES

```
source
1. .138627
  /organism="homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="11"
  /map="11q13"
  /clone="CMB9-22K5"
1. .20726
  /note="assembly_fragment"
20827. 34666
  /note="assembly_fragment"
34767. 46063
  /note="assembly_fragment"
```

```
misc_feature 46166..53516 /note="assembly-fragment"  
misc_feature 53617..61065 /note="assembly-fragment"  
misc_feature 61166..67998 /note="assembly-fragment"  
misc_feature 68099..74526 /note="assembly-fragment"  
misc_feature 74627..81289 /note="assembly-fragment"  
misc_feature 81400..87316 /note="assembly-fragment"  
misc_feature 87417..91541 /note="assembly-fragment"  
misc_feature 91642..95731 /note="assembly-fragment"  
misc_feature 95832..101128 /note="assembly-fragment"  
misc_feature 101229..104948 /note="assembly-fragment"  
misc_feature 105049..110306 /note="assembly-fragment"  
misc_feature 110407..112963 /note="assembly-fragment"  
misc_feature 113070..116484 /note="assembly-fragment"  
misc_feature 116585..119599 /note="assembly-fragment"  
misc_feature 119700..121881 /note="assembly-fragment"  
misc_feature 121982..122703 /note="assembly-fragment"  
misc_feature 122804..125566 /note="assembly-fragment"  
misc_feature 125667..126845 /note="assembly-fragment"  
misc_feature 126946..129349 /note="assembly-fragment"  
misc_feature 129450..131268 /note="assembly-fragment"  
misc_feature 131369..132565 /note="assembly-fragment"  
misc_feature 132666..134031 /note="assembly-fragment"
```

Query Match	21.98;	Score 302.8;	DB 2;	Length 138627;
Best Local Similarity	55.08;	Pred. No. 2e-67;		
Matches 661; Conservative	0;	Mismatches 532;	Indels 9;	Gaps 3;

QY	37	ATTGGCATCAAGGCTTTGACGAAACATCGTGGCGGACAACTGGCGCCACAGGGGATGAGAGG	96
Db	17322	ATTGGCAATTCAGAGCCCTGGCCCAACTAAATTGCTGATGTGGCCCCCAGTGCCTATCCGGAG	17381
QY	97	CAGAAATTCGAGACACTACTCGGCGCGCAAAATGCCGTCGAGCCGACATGACATATAC	156
Db	17382	AATGCATATCAGAGACTACTTTGGCCCTAAGAGTGCCCATTTGATGACTATTTAT	17441
QY	157	CAGTTCCTGATTTAGTATGGAAAGACAGCATGAGAACTGCACAAATGAAAGCTGGTAA	216
Db	17442	CAGTTTCCTGATTTGCTGTCCGAG---GGTGGGGATGTGCTGACGAATGAGAGGGTGA	17498
QY	217	GTCACCTACTATTTCGAAGGAATGTTCAACCGGACAAATTAAGATTACTGGAAAGCGGGATTC	276
Db	17499	ACCACACACCCTGATGAGGCAATGTTCTACCGCACCATTTGCTCATGATGAGGAAGCGATC	17558
QY	277	AAGCCAGTTTATGTGTTTGATGGCAAGCCTCTCGATATGAAGAAACAAGACTTGTCTAA	336
Db	17559	AAGCCGCTGATGTCTTTGATGGCAAGCGCCACAGCTCAAGTCAGGCGAGACTGGCCAA	17618
QY	337	AGATATCTCAAAAAGATGTTGTCGAACCAAGATGTTGACTGAGGGCAGTAAAGAGATGAGAT	396
Db	17619	CGCACTGAGCGCGGCGCTGAGGCGAGGAACACGCTGCGACAGGCTCAAGCTGCTGGGGCC	17678
QY	397	AAAGATGCGATTGAAAATTTGAGCAAGAGGACTGTAAAGTGTACAGGCAACACACAGAA	456

```

Db 17679 GAGCAGAGAGTGGAAAAATTCACCTAGACCGGCTGGTGAAGGTACTAAGACAGCATATAT 17738
QY 457 GATTGTAACGGCTATTAAAGCTTAATGGGGGCTTCCTGTTGTAAGAGCAGCTTCTGAACAA 516
Db 17739 GAGGCAAAACATCTGCTGAGACCTCATGAGGATCCCTTATCTGATGACACCCAGTGAAGCA 17798
QY 517 GAGGCAAAATGTCGAGCCCTTTGCTGAATAAAGCAAGTGTGGGCTGCTGTTGCTGCAAGAAAT 576
Db 17799 GAGGCAAGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 17858
QY 577 ATGAGCTCCCTTACTTATTGGGGCTCCAGCGTTCCTTCTGATTAATGATCAAGTTC 636
Db 17859 ATGAGCTCCCTTACTTATTGGGGCTCCAGCGTTCCTTCTGATTAATGATCAAGTTC 17918
QY 637 AAGAAATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 696
Db 17919 AAAAAAGCTGCAATCCAGAGATTCACCTGAGCGGATTCCTGAGAGAGCTGGGCTGAGC 17978
QY 697 ATGAGCAGATTCATTTGATTTGTCATCTGTGTGATGATGATGATGATGATGATGATGATGAT 756
Db 17979 CAGGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 18038
QY 757 GATATGGGGGGGCAACAGCTGCTGAACCTTATTTGCTGATGATGATGATGATGATGATGATGAT 816
Db 18039 GATATGGGGGGGCAACAGCTGCTGAACCTTATTTGCTGATGATGATGATGATGATGATGATG 18098
QY 817 TTGAGAGATCTTAAATAAAGACAGATATCAAAATCTGAGAGCTGGCTTACCAGAGAGCT 876
Db 18099 GTGGGGGAGCTTGACCCCAACAGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18158
QY 877 CGAGCTGTTGTCAGAGAGCTTAATG---TCACATTTGATATTCCTGAGCTTAATGAGAGCT 933
Db 18159 CACCAAGCTCTTCTTGGAGCTGAGGCTGAGGAGCCAGAGCTCTGAGAGCTGAAGTGAAGC 18218
QY 934 GCACCTGATGAGAGAGCTCTCATTAAGTTCTGTTAAAGATTAATGTTTCAACAGAGAGT 993
Db 18219 GAGCAATATAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 18278
QY 994 CGGGTGAACAAGAGCATATGAGAGATCAAAATCTGCCAAGATTAATTCGTCGCAAGAGAGA 1053
Db 18279 CGAATCCGCGAGTGGGGTCAAGAGAGCTGAGTGAAGAGCCGCAAGAGAGAGAGAGAGAG 18338
QY 1054 CTGAGATCTTTTTCAGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
Db 18339 CTGAGATCTTTTTCAGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18398
QY 1111 ACTTCGATTAACAACAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1170
Db 18399 CCGAAGGAGATCCACTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 18458
QY 1171 AATATATCTTGAATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1230
Db 18459 AATATATCTTGAATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 18518
QY 1231 TT 1232
Db 18519 TT 18520

```

```

RESULT 15
AP002380/c 169053 bp DNA linear HTG 31-MAY-2000
LOCUS Homo sapiens chromosome 11 clone RP11-467L20 map 11q, WORKING DRAFT
DEFINITION SPOUNCE: 35 unordered pieces.
ACCESSION AP002380
VERSION AP002380.1 GI:8131644
SOURCE HTG: HTGS_PHASE1; HTGS_DRAFT.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 169053)

```

# AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong, Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
Homo sapiens 169,053 genomic DNA of 11q  
Published Only in Database (2000) In press  
2 (bases 1 to 169053)  
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong, Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
Direct Submission  
Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8535, Japan (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923, Fax: 81-42-778-9924)

----- Genome Center  
Center: RIKEN Genomic Sciences Center(GSC)  
Center code: RIKEN  
Web site: http://hgp.gsc.riken.go.jp/  
Contact: hattori@gsc.riken.go.jp  
----- Project Information  
Center project name: HumDrafl1  
Center clone name: RP11-467L20  
----- Summary Statistics  
Sequencing vector: PCR products; 100% of reads  
Chemistry: Dye-terminator ET-amersham; 100% of reads  
Assembly program: Phrap, version 0.990329  
Consensus quality: 147827 bases at least Q40  
Consensus quality: 157936 bases at least Q30  
Consensus quality: 162410 bases at least Q20  
Insert size: 165653; sum-of-contigs  
Quality coverage: 4.43x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 35 contigs. The true order of the pieces is not known and the order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1	21322	contig of	21322	bp in length
21423	40245	contig of	18823	bp in length
40346	50492	contig of	10147	bp in length
50593	58569	contig of	7977	bp in length
58670	67385	contig of	8716	bp in length
67486	75248	contig of	7763	bp in length
75349	84453	contig of	9105	bp in length
84554	89492	contig of	4939	bp in length
89593	96445	contig of	6853	bp in length
96546	102631	contig of	6066	bp in length
102732	107199	contig of	4468	bp in length
107300	111820	contig of	4521	bp in length
111921	117068	contig of	5148	bp in length
117169	120901	contig of	3733	bp in length
121002	124072	contig of	3071	bp in length
124173	127611	contig of	3439	bp in length
127712	129401	contig of	1690	bp in length
129502	133141	contig of	3640	bp in length
133242	136380	contig of	3139	bp in length
136481	139425	contig of	2945	bp in length
139526	143281	contig of	3756	bp in length
143382	146449	contig of	3068	bp in length
146550	148828	contig of	2279	bp in length
148929	151045	contig of	2117	bp in length
151146	153372	contig of	2227	bp in length
153473	154996	contig of	1524	bp in length
155097	156763	contig of	1667	bp in length
156864	159132	contig of	2269	bp in length
159233	161349	contig of	2117	bp in length
161450	162850	contig of	1401	bp in length
162951	164533	contig of	1563	bp in length
164634	165829	contig of	1166	bp in length
165930	167130	contig of	1201	bp in length
167231	168290	contig of	1060	bp in length



Db 3399 GAGCAGGAGGTGGAAAAATCTAAGCGGCTGGTAGGTCCTAAGCACACATGAT 3340  
QY 457 GATTGTAACGGCTATTAAAGCTTAAGGAGGATTCCTGTGTAGAGGACCTTCTGAAGCA 516  
Db 3339 GAGGTGAACATCTGCTGAGCCTCAVGGGCATCCCTATCTGTATGCACCCAGTAGGCA 3280  
QY 517 GAAGCAGATGTGACGCCCTTGCAATAACGATAGGTGTGCTGTTCCTTCGAAGAT 576  
Db 3279 GAGGCGAGCTGTGCTGCTGTGAGGCTGCAAGCTTATGCTGGGCTACCGAGAGC 3220  
QY 577 ATGAGCTCCCTTACTTTTGGGCTCCACGGTCTCTTCGTCAATTAAATGATCAAGTTCC 636  
Db 3219 ATGAGCTGCTCACTTGGCAGCCCTGTGCTAATGCGACACTGACCTGACATGAAGCC 3160  
QY 637 AAGAAATACCTGTGATGATTTGATGTGCGAAGSTTTGGAGGACCTGACCTCAC 696  
Db 3159 AAAAAGCTGCAATCGAAGATTCACCTGAGCCGGAATTTGCAAGAGCTGGGCTGAC 3100  
QY 697 ATGAGCAGTTATGATTTGTGATTCCTGTGTGATGTGATTTGATGATCAAAA 756  
Db 3099 CAGGACAGTTGTGTGATCTGTGATCTGTGATGAGTACTGTGATGATATCCGG 3040  
QY 757 GGTATGGGGGGCAACACCTGAACTTATGTCACATGAGTCCATGAAAGCATC 816  
Db 3039 GGTATGGGCCCCAAGGGGCTGTGAGCTCATCCAGAAACACAGAGCATCGAGAGATC 2980  
QY 817 TTGAGAAATCTTAATTAAGACAGATATCAAAATTCCTGAGGACTGGCCTTACCAAGAAGCT 876  
Db 2979 GTGCGGCGACTTGACCCCAACAAAGTACCCTGTGCCAGAAATGGCTCCACAGAGAGCT 2920  
QY 877 CGACGCTTGTCAAGAGGCTTAATG---TCACATGGAATTTCTGAGCTAAATGGACT 933  
Db 2919 CACCAGCTCTTCTTGAGACCTGAGGTGCTGAGCCAGAGCTGTGAGAGCTGAGTGGAGC 2860  
QY 934 GCACCTGATGAGGAGGCTCATAGTTCTGTAAAAAGATATATGTTCAACGAAGAT 993  
Db 2859 GAGCCAAATGAAGAAAGAGCTGATCAAGTTCAATGTGTGAAAAACAGTCTCTGAGAG 2800  
QY 994 CGGGTGAACAAGGCGATTAAGAAATCAAAATCTGCCAAGAAATATCTGCAAGGAAGA 1053  
Db 2799 CGAATCCGCGAGTGGGCTCAAGAGGCTGATAGAGCCGCCAAGGACGACACCAGGGCCGC 2740  
QY 1054 CTCGAGTCTTTTCAAGCCAACTGCCAC--CAGATCAGACGCGCTAAAAAGGAGAG 1110  
Db 2739 CTGGATGATTTCTTCAAGGTGACCGGCTCCTCTTCACTTAAGCGCAAGAGGCAGAA 2680  
QY 1111 ACTTCGATTAACACAGCAAGCGAGCTGCAACAAAGAAACAAAGCTGTGGAAGAAG 1170  
Db 2679 CCCAAGGATCCACTAAGAAAGAAAGACTGGGCGCAGGGAAGTTTAAAAAGGGA 2620  
QY 1171 AAATATCTTGATGCTTGTACAACTAGACATAGCAAGAGCAGCGGTGGCTGATCAC 1230  
Db 2619 AAATTAATGTGTTCCTCCCAATTAATACCTCTTCAACCCAGAAATATTTGCCCTCTGTACCC 2560  
QY 1231 TT 1232  
Db 2559 TT 2558

Search completed: November 5, 2002, 12:54:45  
Job time : 2309.69 secs





FT CDS 85..1224  
FT /\*tag= a  
PM WO200036109-A1.  
XX 22-JUN-2000.  
XX 16-NOV-1999; 99WO-US27147.  
XX PR 15-DEC-1998; 98US-0112332.  
XX (PION-) PIONEER HI-BRED INT INC.  
XX Mahajan PB;  
XX WPI: 2000-452026/39.  
XX DR P-PSDB; AAY95307.  
XX  
XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA  
PT recombination and repair in transgenic plants, e.g. for gene targeting  
PT and the production of male sterile plants -  
XX  
XX Claim 1; Page 69-71; 85pp; English.  
XX  
XX The present sequence is that of maize cDNA coding for RAD2/FEN-1  
CC (see AAY95307). The corresponding RNA was isolated from immature  
CC ear tissue from 2 ears of a B73 maize line. Rad2/FEN-1 is a  
CC structure specific endonuclease which under certain conditions also  
CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to  
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.  
CC The protein is involved in the regulation of DNA repair and  
CC recombination in plant systems and therefore may be used for  
CC improving gene targeting during further recombinant DNA protocols  
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential  
CC in DNA replication and nucleotide excision and repair reactions.  
CC The exolytic activity is involved in double strand break repair and  
CC end joining. The protein is also useful in strand exchange  
CC reactions during homologous recombination. These functions may be  
CC useful in gene targeting and in the production of male sterile  
CC plants. The efficacy of gene targeting can be improved by the  
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
CC be produced by the down regulation of Rad2/FEN-1 expression.  
XX  
XX Sequence 1463 BP; 466 A; 292 C; 361 G; 344 T; 0 other;  
SQ  
Query Match 97.8%; Score 1350.2; DB 21; Length 1463;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1355; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 GCACCCAGCGCTCGCGCCACAGCGCGCGAGAGATGGGATCAAGGTTTGAGCAAA 60  
DB 49 CGCGCCCGCCACCGCCACAGCGCGCGAGATGGGATCAAGGTTTGAGCAAA 108  
QY 61 CTGCTGGCGGCAATGCGCCCAAGGCGATGAAGAGCAGAAAGTTGAGAGCTACTTCGGC 120  
DB 109 CTGCTGGCGGCAATGCGCCCAAGGCGATGAAGAGCAGAAAGTTGAGAGCTACTTCGGC 168  
QY 121 CGCAAAATCGCGCTCGAGCCAGCATGAGCATATACCAAGTTCCGATTTAGTTGAGAG 180  
DB 169 CGCAAAATCGCGCTCGAGCCAGCATGAGCATATACCAAGTTCCGATTTAGTTGAGAG 228  
QY 181 ACAGGATGGAAGCTCTCACAATAGAGCTGGTGAAGTCACTAGTCATTTGCAAGGATG 240  
DB 229 ACAGGATGGAAGCTCTCACAATAGAGCTGGTGAAGTCACTAGTCATTTGCAAGGATG 288  
QY 241 TTCAACCGGACAAATTAAGTTTACTGGAAGCGGAGATCAAGCCAGTTATGTTTGGATGCG 300  
DB 289 TTCAACCGGACAAATTAAGTTTACTGGAAGCGGAGATCAAGCCAGTTATGTTTGGATGCG 348  
QY 301 AAGCCTCTGATATGAGAACAAGAGCTTGCTAAAGATTAAGTCAAAAGAGATGATGCA 360  
DB 349 AAGCCTCTGATATGAGAACAAGAGCTTGCTAAAGATTAAGTCAAAAGAGATGATGCA 408

QY 361 ACCAAGATCTGACTGAGCAGTAGAGGTAGAGATTAAGATGCGATTAAGAAAATTTGAGC 420  
DB 409 ACCAAGATCTGACTGAGCAGTAGAGGTAGAGATTAAGATGCGATTAAGAAAATTTGAGC 468  
QY 421 AAGAGACTGTAAAGGTCACAGAGCAACACAGCAAGATTTTAACGGCTATTAGACTT 480  
DB 469 AAGAGACTGTAAAGGTCACAGAGCAACACAGCAAGATTTTAACGGCTATTAGACTT 528  
QY 481 ATGGGGGTTCTGTTGTAGAGGACCTTGTGAAGCGAAGACAGATGTGCAGCCCTTTCG 540  
DB 529 ATGGGGGTTCTGTTGTAGAGGACCTTGTGAAGCGAAGACAGATGTGCAGCCCTTTCG 588  
QY 541 ATAAACGATTAAGTGTTCGCTGTTCGTCAGAAAGATATGAGCTCCCTACTTTGGGCT 600  
DB 589 ATAAACGATTAAGTGTTCGCTGTTCGTCAGAAAGATATGAGCTCCCTACTTTGGGCT 648  
QY 601 CCAGGTTCCCTGCTCAATTAATGATCCAGGTTCACAGAAATATCCTGTGATGATTT 660  
DB 649 CCAGGTTCCCTGCTCAATTAATGATCCAGGTTCACAGAAATATCCTGTGATGATTT 708  
QY 661 GATGTGCCAAGGTTTGTGAGAGGCTTGAACATCAACATGAGCAGTTCAATGATTTGTC 720  
DB 709 GATGTGCCAAGGTTTGTGAGAGGCTTGAACATCAACATGAGCAGTTCAATGATTTGTC 768  
QY 721 ATCTGTGTGATGTGACTATTTGATAGCATCAAGGTATCGGGGGCAACAGCTCTG 780  
DB 769 ATCTGTGTGATGTGACTATTTGATAGCATCAAGGTATCGGGGGCAACAGCTCTG 828  
QY 781 AACTTATTCGTCACATGGGCTCATGAAAGCATCTTGGAGATCTTAATAAGACAGA 840  
DB 829 AACTTATTCGTCACATGGGCTCATGAAAGCATCTTGGAGATCTTAAATAAGACAGA 888  
QY 841 TATCAAAATTCGAGGACTGGGCTTACCAAGAGCTGAGCTGTTCGCAAGGAGCTTAAT 900  
DB 889 TATCAAAATTCGAGGACTGGGCTTACCAAGAGCTGAGCTGTTCGCAAGGAGCTTAAT 948  
QY 901 GTCACTTGTGATATTCCTGAGCTTAAGTGAAGTGCACCTGATGAGAGGGCTCATAGT 960  
DB 949 GTCACTTGTGATATTCCTGAGCTTAAGTGAAGTGCACCTGATGAGAGGGCTCATAGT 1008  
QY 961 TTCCGTGTAAGAAATATGTTTCAACGAAGTGGGTGCAAGGCAATGAGAGATGTC 1020  
DB 1009 TTCCGTGTAAGAAATATGTTTCAACGAAGTGGGTGCAAGGCAATGAGAGATGTC 1068  
QY 1021 AAATCTGCAAGAAATTAATGCTGCAAGGAAGACTGAGTCTTTTCAAGCCACTGTC 1080  
DB 1069 AAATCTGCAAGAAATTAATGCTGCAAGGAAGACTGAGTCTTTTCAAGCCACTGTC 1128  
QY 1081 ACCACATGACACCGCTTAAGGGAAGAGAGCTGGAATTAACCAAGCAAGGACGCTGGC 1140  
DB 1129 ACCACATGACACCGCTTAAGGGAAGAGAGCTGGAATTAACCAAGCAAGGACGCTGGC 1188  
QY 1141 AACAGAAACAAAGGCTGTGGAAGAAAGAAATTAATCTGTGATGCTGTGATACACTA 1200  
DB 1189 AACAGAAACAAAGGCTGTGGAAGAAAGAAATTAATCTGTGATGCTGTGATACACTA 1248  
QY 1201 CGACTAGCAAGGCGGTGGGATCCTGCTAGATTTTAATACCCCTGTGTTTA 1260  
DB 1249 CGACTAGCAAGGCGGTGGGATCCTGCTAGATTTTAATACCCCTGTGTTTA 1308  
QY 1261 ACTCAGAGCTTTGTTAAAGTTTCGCCCATGTTTCAAGCTGGGTAAGTTAGTGTGTTG 1320  
DB 1309 ACTCAGAGCTTTGTTAAAGTTTCGCCCATGTTTCAAGCTGGGTAAGTTAGTGTGTTG 1368  
QY 1321 AAGAGATTGGTGTACCAAGTACCAAACTTATGCTGTTT 1363  
DB 1369 AAGAGATTGGTGTACCAAGTACCAAACTTATGCTGTTT 1411  
RESULT 3  
AAA27924  
ID AAA27924 standard; cDNA; 1541 BP.  
XX

AC	AAA27924.
XX	12-SEP-2000 (first entry)
DT	
XX	
DE	Maize Rad2/FEN-1 cDNA.
XX	
KW	Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KM	endonuclease; exonuclease; DNA repair; gene targeting; ss.
XX	
OS	Zea mays.
XX	
XX	
FH	Location/Qualifiers
FT	Key 79..1218
FT	CDS /*tag= a
XX	
PN	WO200036109-A1.
PD	
XX	22-JUN-2000.
PE	
PP	16-NOV-1999; 99MO-US27147.
XX	
PR	15-DEC-1998; 98US-0112332.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
PX	
PI	Mahajan PB.
XX	
DR	WPI: 2000-452026/39.
PT	P-PDB; AAY95308.
PT	
PT	Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA recombination and repair in transgenic plants, e.g. for gene targeting and the production of male sterile plants -
XX	
PS	Example 1; page 73-74; 85pp; English.
XX	
CC	The present sequence is that of maize cDNA (ATCC PTA-533) coding for RAD2/FEN-1 (see AAY95308). The corresponding RNA was isolated from a B73 line seedling after a 10 day drought, heat shocked for 10 hr, and allowed to recover under normal conditions. Rad2/FEN-1 is a structure specific endonuclease which under certain conditions also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to produce the Rad2/FEN-1 polypeptides in transgenic plant cells. CC The protein is involved in the regulation of DNA repair and recombination in plant systems and therefore may be used for improving gene targeting during further recombinant DNA protocols involving plants. RAD2/FEN-1 endonucleolytic activity is essential in the DNA replication and nucleotide excision and repair reactions. CC The exolytic activity is involved in double strand break repair and end joining. The protein is also useful in strand exchange reactions during homologous recombination. These functions may be useful in gene targeting and in the production of male sterile plants. The efficacy of gene targeting can be improved by the overexpression of exogenous Rad2/FEN-1 while male sterile plants can be produced by the down regulation of Rad2/FEN-1 expression.
XX	
XX	
SQ	Sequence 1541 BP; 473 A; 308 C; 377 G; 383 T; 0 other:
	Query Match 97.7%; Score 1348.6; DB 21; Length 1541;
	Best Local Similarity 99.3%; Pred. No. 0;
	Matches 1354; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
OY	1 CGAGCCACGGGTGCGGCCAAGCGCGGCAGACAGCATGCGCATCAAGSGTTTGACGA 60
DB	43 CGCGGCCCGCCACCAGCCGCAAGCGCGCAGACAGAATGCGCATCAAGSGTTTGACGA 102
OY	61 CTCTGTGGCGGCAATGCGGCCAAGGCGATGAAGAGACAGAAAGTTCGAGAGCTACTTGCGC 120
DB	103 CTGCTGGCGGACAATGCGGCCCAAGGGGATGAAGAGACAGAAAGTTCGAGAGCTACTTGCGC 162
OY	121 CGCAAAATGCGCGTGCAGCGCCAGCATGAGCATATTACCAGTTCTGATTTAGTTGGAAG 180
DB	163 CGCAAATGCGCGTGCAGCGCCAGCATGAGCATATTACCAGTTCTGATTTAGTTGGAAG 222

Qy	181	ACAGCGATGGAACTCTCCAAATGAAGCTGGTGAAGTCACTAGTCAATTGGCAAGAAATG	240
Db	223	ACAGCGATGGAACTCTCCAAATGAAGCTGGTGAAGTCACTAGTCAATTGGCAAGAAATG	282
Qy	241	TTTCAACCGGACATTAAGATTACTCTGAAGCGGGAAATCAAGCCAGTTATTTGTTTGAATGC	300
Db	283	TTTCAACCGGACATTAAGATTACTCTGAAGCGGGAAATCAAGCCAGTTATTTGTTTGAATGC	342
Qy	301	AAGCTCTCTATATGAAGAACCAAGAGCTTGGTAAAGATACTCAAAAAGATGATGCA	360
Db	343	AAGCTCTCTATATGAAGAACCAAGAGCTTGGTAAAGATACTCAAAAAGATGATGCA	402
Qy	361	ACCAAAGATCTGACTGAGCGATGAGAGTGAAGATTAAGATGGGATTGAAAAATTAGC	420
Db	403	ACCAAAGATCTGACTGAGCGATGAGAGTGAAGATTAAGATGGGATTGAAAAATTAGC	462
Qy	421	AAGGAGACTGTAAAGGTCACAAAGCAACCAACGAAATTTGAAACGGCTATTAGACTT	480
Db	463	AAGGAGACTGTAAAGGTCACAAAGCAACCAACGAAATTTGAAACGGCTATTAGACTT	522
Qy	481	ATGGGGTTTCTGTGTTGAGAGCCACTTCTGAAGCAGAAGATGTGCAAGCCCTTTGC	540
Db	523	ATGGGGTTTCTGTGTTGAGAGCCACTTCTGAAGCAGAAGATGTGCAAGCCCTTTGC	582
Qy	541	ATTAACGATTAAGGTTGCGGTGGTGGTCGAAGATATGATGATCCCTTACTTTGGGGCT	600
Db	583	ATTAACGATTAAGGTTGCGGTGGTGGTGGTCGAAGATATGATGATCCCTTACTTTGGGGCT	642
Qy	601	CCACGGTTCCTTGCTCATTTTAATGATCCAAAGTTCCACAGAAATACCTGTATGGAATTT	660
Db	643	CCACGGTTCCTTGCTCATTTTAATGATCCAAAGTTCCACAGAAATACCTGTATGGAATTT	702
Qy	661	GATGTTGCCAAGTTTGGAGAGGAGCTGAACATCCATGAGACAGTCAATGATTGTTGCG	720
Db	703	GATGTTGCCAAGTTTGGAGAGGAGCTGAACATCCATGAGACAGTCAATGATTGTTGCG	762
Qy	721	ATTCGTGTGATGATGATATTTGATATGATCATTCAAGATATCGGGGGGCAAAAGCTCTG	780
Db	763	ATTCGTGTGATGATGATATTTGATATGATCATTCAAGATATCGGGGGGCAAAAGCTCTG	822
Qy	781	AAACTTATTTCTCAACATGGSTCCATGGAAGATCTTGAGAAATCTTAATAAAGACGA	840
Db	823	AAACTTATTTCTCAACATGGSTCCATGGAAGATCTTGAGAAATCTTAATAAAGACGA	882
Qy	841	TATCAATTTCTGAGAGCTGCGCTTACCAGAAGCTCGAGCGTTGTTCAAGAGCCTAAT	900
Db	883	TATCAATTTCTGAGAGCTGCGCTTACCAGAAGCTCGAGCGTTGTTCAAGAGCCTAAT	942
Qy	901	GTTCACATTGATATTTCTGAGCTTAAATGACATGCACTGATGAGAGGCTCTTAATGT	960
Db	943	GTTCACATTGATATTTCTGAGCTTAAATGACATGCACTGATGAGAGGCTCTTAATGT	1002
Qy	961	TTTCCGTGTTAAAGATATATGTTTCAACGAAGATCGGGTGAACAAAGGCCATTAGAAAGATC	1020
Db	1003	TTTCCGTGTTAAAGATATATGTTTCAACGAAGATCGGGTGAACAAAGGCCATTAGAAAGATC	1062
Qy	1021	AAATCTGCCAAGATTAATTCGTCCCAAGAAAGTCTGATCTTTTTCAGAGCAATCTCC	1080
Db	1063	AAATCTGCCAAGATTAATTCGTCCCAAGAAAGTCTGATCTTTTTCAGAGCAATCTCC	1122
Qy	1081	ACCAATCTGACGAGCTTAAACGGAAGAGCATTCGGATTAACCAAGCAAGGCAAGGCACTCTGC	1140
Db	1123	ACCAATCTGACGAGCTTAAACGGAAGAGCATTCGGATTAACCAAGCAAGGCAAGGCACTCTGC	1182
Qy	1141	AACCAAAACCAAGAGCTGTGTGGAAGAAAGCAATTAATCTTGATGCTTGATGTACACTA	1200
Db	1183	AACCAAAACCAAGAGCTGTGTGGAAGAAAGCAATTAATCTTGATGCTTGATGTACACTA	1242
Qy	1201	CGAGCTACGAAAGACAGCGGTGGTGGTATCATCTCCCTAGATTATTTTAACCTCCCTGTTTA	1260
Db	1243	CGAGCTACGAAAGACAGCGGTGGTGGTATCATCTCCCTAGATTATTTTAACCTCCCTGTTTA	1302



QY 1261 ACTCAGAGCTTTGGTAAAGTTCCGCCATGTTTCAGCTGGGGTAAGTACTGTTGTTG 1320  
Db 1303 ACTCAGAGCTTTGGTAAAGTTGCTCATGTTTCAGCTGGGGTAAGTACTGTTGTTG 1362  
QY 1321 AAGAGATTGTTGTTACCAAGTAAACAACCTATCGCTGTTT 1363  
Db 1363 AAGAGATTGTTGTTACCAAGTAAACAACCTATCGCTGTTT 1405

RESULT 4  
AAZ27926  
ID AAAZ27926 standard; cDNA, 1478 BP.  
XX  
AC AAAZ27926;  
XX  
DT 12-SEP-2000 (first entry)  
XX  
DE Maize Rad2/FEN-1 cDNA.  
XX  
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;  
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 97..1236  
FT /tag= a  
XX  
PN WO200036109-A1.  
XX  
PD 22-JUN-2000.  
XX  
PF 16-NOV-1999; 99WO-US27147.  
XX  
PR 15-DEC-1998; 98US-0112332.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Mahajan PB;  
XX  
DR WPI; 2000-452026/39.  
DR P-PSDB; AAY95310.  
XX  
PT Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA  
PT recombination and repair in transgenic plants, e.g. for gene targeting  
XX  
PS Example 1; Page 77-79; 85pp; English.  
XX  
CC The present sequence is that of maize cDNA coding for RAD2/FEN-1  
CC (see AAY95310). The cDNA was isolated from a library prepared  
CC from B73 line seed vitreous endosperm RNA. Rad2/FEN-1 is a  
CC structure specific endonuclease which under certain conditions also  
CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to  
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.  
CC The protein is involved in the regulation of DNA repair and  
CC recombination in plant systems and therefore may be used for  
CC improving gene targeting during further recombinant DNA protocols  
CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential  
CC in DNA replication and nucleotide excision and repair reactions.  
CC The exolytic activity is involved in double strand break repair and  
CC end joining. The protein is also useful in strand exchange  
CC reactions during homologous recombination. These functions may be  
CC useful in gene targeting and in the production of male sterile  
CC plants. The efficacy of gene targeting can be improved by the  
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
CC be produced by the down regulation of Rad2/FEN-1 expression.  
XX  
SQ Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other;

Query Match 96.6%; Score 1334.2; DB 21; Length 1478;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1345; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 CGACCCAGCGCTCCGGCCACAGCCGCCGACAGATGGCATCAAGGTTTGACGAA 60  
Db 61 CGCGCCCGCCACCCGCCACAGCGCCGCCGACAGATGGCATCAAGGTTTGACGAA 120  
QY 61 CTGCTGGCGCAATGCGCCCAAGGGGATGAAGAGACAGAGTTCGAGAGCTTCGCG 120  
Db 121 CTGCTGGCGCAATGCGCCCAAGGGGATGAAGAGACAGAGTTCGAGAGCTTCGCG 180  
QY 121 CGCAAAATCGCCGTCGACGCCAGCATGACATATACCACTTCTGATTTGGAAG 180  
Db 181 CGCAAAATCGCCGTCGACGCCAGCATGACATATACCACTTCTGATTTGGAAG 240  
QY 181 ACAGCATGAAATCTCAGCAAAATGAAGCTGTGAAGTCACTAGTCATTTGCAAGAA 240  
Db 241 ACAGCATGAAATCTCAGCAAAATGAAGCTGTGAAGTCACTAGTCATTTGCAAGAA 300  
QY 241 TTCAACCGGCAATATAGATTACTGGAAGCGGGAATCAAGCCATTATTTGATG 300  
Db 301 TTCAACCGGCAATATAGATTACTGGAAGCGGGAATCAAGCCATTATTTGATG 360  
QY 301 AAGCTCCTGATATGAAGAAACAAGAGCTTGTAAAGATCTCAAAAAGAGATGCA 360  
Db 361 AAGCTCCTGATATGAAGAAACAAGAGCTTGTAAAGATCTCAAAAAGAGATGCA 420  
QY 361 ACCAAAGATCTGACGTAGCGAGGTAGAGTACGATTAAGATCGCTTGAATAAT 420  
Db 421 ACCAAAGATCTGACGTAGCGAGGTAGAGTACGATTAAGATCGCTTGAATAAT 480  
QY 421 AAGAGACTGTAAGGTACACAGGCAACACAGAGATTGTAAACGGCTATTAGACT 480  
Db 481 AAGAGACTGTAAGGTACACAGGCAACACAGAGATTGTAAACGGCTATTAGACT 540  
QY 481 ATGGGGTTCCTGTTGTAAGAGGACCTTCGAGGCAAGCAAGATGTCAGCCCTT 540  
Db 541 ATGGGGTTCCTGTTGTAAGAGGACCTTCGAGGCAAGCAAGATGTCAGCCCTT 600  
QY 541 ATAAAGATTAAGGTGTTGCTGCTTCAAGAGATTAAGGACTCCCTTACTTTGG 600  
Db 601 ATAAAGATTAAGGTGTTGCTGCTTCAAGAGATTAAGGACTCCCTTACTTTGG 660  
QY 601 CCAGGTTCTCTGCTGATTTATGATGATCCAAAGTTCAGAAATATCTGATG 660  
Db 661 CCAGGTTCTCTGCTGATTTATGATGATCCAAAGTTCAGAAATATCTGATG 720  
QY 661 GATGTTGCCAAGGTTTGGAGAGCTTGAATCAATGACAGTTCATTTGATTT 720  
Db 721 GATGTTGCCAAGGTTTGGAGAGCTTGAATCAATGACAGTTCATTTGATTT 780  
QY 721 ATCCGTGTGATGATGATTTGATGATGATCAAAAGTTCGGGGGCAAAAGCT 780  
Db 781 ATCCGTGTGATGATGATTTGATGATGATCAAAAGTTCGGGGGCAAAAGCT 840  
QY 781 AAAGTATTTCGTCATGATGGTTCATGAAGAGCATCTTGGAGATCTTAATAAG 840  
Db 841 AAAGTATTTCGTCATGATGGTTCATGAAGAGCATCTTGGAGATCTTAATAAG 900  
QY 841 TTTCATATTCCTGAGAGCTGACCTTACCAAGAGCTGAGAGCTTTCAGAG 900  
Db 901 TTTCATATTCCTGAGAGCTGACCTTACCAAGAGCTGAGAGCTTTCAGAG 960  
QY 901 GTTCATATTCCTGAGAGCTGACCTTACCAAGAGCTGAGAGCTTTCAGAG 960  
Db 961 GTTCATATTCCTGAGAGCTGACCTTACCAAGAGCTGAGAGCTTTCAGAG 1020  
QY 961 TTCTGTGTAAGATATGTTTCAAGAGATGGGTGCAAGGCAATAGAGAT 1020  
Db 1021 TTCTGTGTAAGATATGTTTCAAGAGATGGGTGCAAGGCAATAGAGAT 1080  
QY 1021 AATCTGCCAAGATATATCTGCGAAGAGAGTCAAGTCTTTTCAAGCAACT 1080  
Db 1081 AATCTGCCAAGATATATCTGCGAAGAGAGTCAAGTCTTTTCAAGCAACT 1140



```

RESULT 6
ID AAX02107 standard: cDNA: 1144 BP.
XX
AC AAX02107;
XX
DT 23-APR-1999 (first entry)
XX
DE Human FEN-1 cDNA.
XX
KM FEN-1: human; flap endonuclease; detection: diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1144
FT /tag="a
FT /product="FEN-1"
XX
XX US5874283-A.
XX
XX 23-FEB-1999.
XX
XX 30-MAY-1995; 95US-0455968.
XX
XX 30-MAY-1995; 95US-0455968.
XX
XX (HARR/) HARRINGTON J J.
XX (HSIE/) HSIEH C.
XX (LIEB/) LIEBER M R.
XX
XX Harrington JJ, Hsieh C, Lieber MR;
XX
XX WPI: 1999-179985/15.
XX P-PSDB; AAM92504.
XX
XX DNA encoding flap endonuclease polypeptides - useful for producing
XX e.g. recombinant polypeptides
XX
XX Disclosure: Fig 1B; 58pp: English.
XX
XX This sequence encodes a human FEN-1 (flap endonuclease) protein. This
XX protein can be used in methods for detecting a pathological condition in
XX a patient, for diagnostic purposes, for screening for antineoplastic
XX agents and carcinogens, for diagnostic staging of neoplasia, for
XX producing recombinant flap endonuclease for use as research or
XX diagnostic reagents, for producing transgenic nonhuman animals expressing the
XX novel polypeptides encoded by a transgene. The invention also provides
XX a novel molecular cloning techniques and reagents involving cleavage of
XX a flap or nick with a flap endonuclease.
XX
XX Sequence 1144 BP; 292 A; 283 C; 353 G; 216 T; 0 other;
XX
XX Query Match 21.9%; Score 302.6; DB 20; Length 1144;
XX Best Local Similarity 56.6%; Pred. No. 1.5e-77;
XX Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

```

```

DB 178 ACCACAGCCACCTGATGGCATGTTCTACCGCACCATTCGATGATGGAAGAGCATC 237
XX
XX 277 AAGCCAGTTATGTTTGTATGGCAAGCCTCCTGATATGAAGAAACAAGCCTTGCTAA 336
XX
DB 238 AAGCCGCTATGCTCTTGTATGGCAAGCCTCCTGATGATGATGATGATGATGATG 297
XX
XX 337 AGTACTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGGAGTATGATGATGAT 396
XX
DB 298 CCGAGTGAAGCGCGGAGCTGAGCAGAGAAGCAGCTGACGAGGCTCAGGCTCGTGGG 357
XX
XX 397 AAGATGCCATGTAATAATTTGAGCAAGAGACTGTAAGGTGACAAAGCAGCAGCA 436
XX
DB 358 GAGCAGAGAGTGGTGAATAATTCATGAGCGCTGGTGAAGTCTAGTAAAGCAATGAT 417
XX
XX 457 GATTGTAAGCGCTATTAAGACTATGAGGAGGCTCTGTTGAGAGCAGCTTCGAGCA 516
XX
DB 418 GAGTGCMAACATCTGTGAGCTCATGGGCAATCCCTTATCTTGATGACCCAGTGA 477
XX
XX 517 GAGCGAATATGACGCCCTTTGCTAATACGATAGGTGTTGCTGTTGCTGAGAGAT 576
XX
DB 478 GAGGCCAGCTGTGCTGCCCTGAGAGCTGCAAGTCTATGCTGGGCTACCGAGGAC 537
XX
XX 577 ATGAGCTCCCTTACTTGTGAGGCTCCAGGCTTCTCTGCTCATTTAATGATCCAA 636
XX
DB 538 ATGAGCTGCTCACCCTGCGGAGCCTCTGTGCAATGCGACACGACGACGACGAG 597
XX
XX 637 AAGAAATATCCTGTGATGATGATTTGATGTTGCAAGCTTTGAGAGAGCTTGA 636
XX
DB 598 AAAAAGCTCCCATCAGGAATTCACCTGAGACCCGATTCGAGAGAGCTGGGCTG 657
XX
XX 697 ATGAGCACTTATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 756
XX
DB 658 CAGGAGCACTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
XX
XX 757 GGTATCGGGGGGCAAGAGCTGTGAACCTATTCGTCAACATGGGTGATGAAAGAT 816
XX
DB 718 GGTATGAGGCGCCCAAGGGGCTGTGACCTCATCCAGAGCAAGAGAGATGAGAG 777
XX
XX 817 TTGAGAAATCTTAATTAAGACAGATATCAAAATTCCTGAGAGCTGAGCTTAC 876
XX
DB 778 GTGCGGCGACTTGACCCCAACAAATGATACCTGTGCGCAAGAAATTTGGTCC 837
XX
XX 877 CGAGCTTTGTAAGAGCCTAATG---TCACATTTGATTTCTGACCTTAATGAG 933
XX
DB 838 CACCACTCTTTCTTGAACCTGAGGTGCTGAGCCAGAGCTGTGAGAGCTGAAG 897
XX
XX 934 GCACCTGATGAGAGAGGCTCATAGTTTCCGTGTAAGAAAGATTAATGTTCAAG 993
XX
DB 898 GAGCCAAATGAAGAGAGCTGATCAAGTTCAATGTGTGTAAGAGAGCTTCTGAG 957
XX
XX 994 CGGGTACAAAGGCCATGAGAGATCAAAATCGCCAAAGAAATTAATGTCGCAAG 1053
XX
DB 958 CGAATCCGACGTGGGTCAAGAGAGCTGAGTAAGAGCCGCCAAGGAGAGAGCC 1017
XX
XX 1054 CTCGAGTCTTTTTCAGGCACTGCCACCATGAGCAGC 1094
XX
DB 1018 CTGAGTATTTCTTCAAGGTGACGGCTCAGCTCTTCAAGC 1058
XX

```

```

RESULT 7
ID AAX02111 standard: DNA: 2033 BP.
XX
AC AAX02111;
XX
DT 23-APR-1999 (first entry)
XX
DE Human FEN-1 genomic DNA.
XX
KM FEN-1: human; flap endonuclease; detection: diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage; ss.
XX

```

OS	Homo sapiens.	Location/Qualifiers
XX	Key	104..1240
EH	CDS	/*tag= a
FT		/product= "FEN-1"
FT		
XX		
PN	US5874283-A.	
XX		
PD	23-FEB-1999.	
XX		
PF	30-MAY-1995;	95US-0455968.
XX		
PR	30-MAY-1995;	95US-0455968.
XX		
PA	(HARR/) HARRINGTON J J.	
PA	(HSIE/) HSIEH C.	
PA	(LIEB/) LIEBER M R.	
XX		
PI	Harrington JJ, Hsieh C, Lieber MR;	
XX		
DR	WPI: 1999-179985/15.	
XX	P-PSDB: AAW92508.	
XX		
PT	DNA encoding flap endonuclease polypeptides - useful for producing e.g. recombinant polypeptides	
XX		
PS	Disclosure: Fig 5A-B; 58pp; English.	
XX		
CC	This sequence encodes a human FEN-1 (flap endonuclease) protein. This protein can be used in methods for detecting a pathological condition in a patient, for diagnostic purposes, for screening for antineoplastic agents and carcinogens, for diagnostic staging of neoplasia, for producing recombinant flap endonuclease for use as research or diagnostic reagents, for producing antibodies reactive with the novel polypeptides, for producing transgenic nonhuman animals expressing the novel polypeptides encoded by a transgene. The invention also provides novel molecular cloning techniques and reagents involving cleavage of a flap or nick with a flap endonuclease.	
CC		
XX	Sequence 2033 BP, 500 A; 493 C; 579 G; 461 T; 0 other:	
XX		
SQ		
	Query Match	20.0%; Score 276.6; DB 20; Length 2033;
	Best Local Similarity	55.9%; Pred. No. 7.4e-70;
	Matches 638; Conservative	0; Mismatches 479; Indels 24; Gaps
OY	37 ATGGGCATCAAGGCTTGACGAACCTGCGCGGACAAATGCCCAAGCGATGAAGGAG	96
DB	104 ATGGAATTTACAGGCCCTTGCCAACTAATGCTGATGTGGCCCCAGTGCATCCGTGAG	163
OY	97 CAGAAAGTTCAGAGACTACTTCGGCCGCAAAATGCGCGTCGACGCCAGCATGACATATAC	156
DB	164 AATACATCATCAAGAGACTCTTTGGTGGCAAAAGTGGCATGCATCGCTTCACATCATC	223
OY	157 CAGTTCTGATTTAGTTGGAAGGACAGCGATGGAACCTTCACAAATGAAGCTGGTGA	216
DB	224 CAGTTCTGATTTGCTGTTCTGCACG--GATGGGATGCTCTGCAGAAACGAGGAGGTGAG	280
OY	217 GTCATGTCATTTGCAAGGAATTTCAACCGGACCAATTAAGATTAATCTGAAGCGGATC	276
DB	281 ACCACACGCC--TGATGGCATTTTATACCGTACCATGCG--CATGAGAAATGGCATC	334
OY	277 AAGCCAGTTTATGTTTGTGATGGCAAGCCTCTGATATGAAGAAACAAGACTTGTCTAA	336
DB	335 AAGCTGTGTACGTCCTTTGATGGCAACACACAGCGTAAGTCAAGCGCTGGCCAAAG	394
OY	337 AGATACCTCAAAAAGATGATGCACCAACAAGATCTGACTGAGGCAAGTAAAGTGAAGAT	396
DB	395 CGCAGTGAAGAGCCGCGGAGGCTGAGAAAGCAACTGACGAGGCTGCACAGAGCTGGATG	454
OY	397 AAAGTGGCATTTGAAAAATTTGAGCAAGAGGACTTAAGATGCACAAGGCAACAAGAA	456
DB	455 GAGGAGGAGGATGGAGAAAGTTCCACCAAGAGGCTGTGTGAAGATGCACCAAGCAACAATGAT	514

QY	457	GAATTGTAACGCGTATTAAAGACTTATGSGGGGTTCCTGTGTAGAGGCGACCTTCTGAAGCA	516
Db	515	GAGTGCACAACCTGCTGAGCCCTCATGGGCATCCCTTACTTGTATGCAACCCAGCGAGCA	574
QY	517	GAACCAACAATGTCACACCCCTTGTGCATTAACGATNAAGTGTTCGCTGTTGGTTCAGAAGAT	576
Db	575	GAGGCCAGCTGTGCTGCTCCCTGGCAAAAGCGTGGCCAAAGATCTATGCTGGCGGCACGAGAGAC	634
QY	577	ATGGACTCCCTTACTTGTGTGGGCGCTCCACGCGTTCCTTGTGTATTTAATGATCCAAAGTTC	636
Db	635	ATGAGCTGCCTCACTTGTGTGGCAGCGCCCGCTGTAAAGCCAGCACTTAACCTGCAGTGAAGCC	694
QY	637	AAGAAATFACCTGMAATGGATTTGATGTGTCAGAGGTTTGGAGAGAGCTTGAATCACC	696
Db	695	AAGAAAGCTGCCCATCCAAAGATTCCATCTGAGCCGCGCTGCGAGAGAGCTGGGTGTAAC	754
QY	697	ATGAGACAGTTCATTGATTTGTGTCATCCGTGTGATGATGTGACTATTTGTATGATCAATCA	756
Db	755	CAGAGAGGAGTTGTGGATCTGTGCATTCCTGTGGTAGCCACTACTGAGAGAGATCCGT	814
QY	757	GGTATCGGGGGGCAACAGCTCTGAAACTTATTTGTCAAACATGGGTCCATGGAAGCATC	816
Db	815	GGCATTTGGCCGCAAGCGGCGCTGTGGATCTCATCCAGAAACATMAAGCATGAGAGATTC	874
QY	817	TTTGAGATCTTAAATTAAAGACAGATATCAAAATTCCTGAGAGACTGGCTTTACCAAGAACT	876
Db	875	GTAAGCGCGCTGGACCCCAAGCAAGTACCCCGTTCCAGAGAACTGGCTCCACAGAGAAAGCC	934
QY	877	CGAGCGTTGTCAGAGAGCCCTAA--TGTACATTGGATATTTCTGAGACTTAAATGTGACT	933
Db	935	CAGCAGCTCTCTCGAGCCAGAAAGTATGAGACCCAGAGTCTGTGGAGCTGAAGTGGAGC	994
QY	934	GCACCTGATGAGAGGGGTCCATTAAGTTCTCGTAAAGATATGTTTACAGGAAGAT	993
Db	995	GAGCCAAATGTAGAAAGAGTGTGTCATAATTTATGTGTGTTGAAAACAGATTTTGTGAAG	1054
QY	994	CGGTTGACAAAGGCCCATAGAAAGATCAAAATCTGCCAAGATTAATCGTGCAGAGAGA	1053
Db	1055	CGAATTTGCCACGTGGGGTTCAGCGGCTGAGTAAGAGCCGCGCAGGCGAGCAACCCAGGGAGCG	1114
QY	1054	CTCGAGTCCCTTTTTCAGGCCAAGCTGCCACACATCAAGCAACCGCTAAACGGAAGAGACT	1113
Db	1115	CTCGATGATTTCTTCAAGGTACAGGCTCACTCTCTCAGC-----TAAAGCG	1162
QY	1114	TGCGATTAAACAGCAGCAGCTGCGAACAAGAAAACAAGGCTGTGGAAAAGAGAAA	1173
Db	1163	AAGGAGCCAGAACCCAGAGGGCTCTGTAAAGAAAGAAAGCAAGATGTGGGAGCGGGGAG	1222
QY	1174	T 1174	
Db	1223	T 1223	
RESULT 8			
AC	ABLI4287	standard; cDNA: 1300 BP.	
XX	XX		
XX	ABLI4287;		
XX	26-MAR-2002	(first entry)	
DE	XX	Drosophila melanogaster expressed polynucleotide SEQ ID NO 37343.	
XX	KW	Drosophila: developmental biology; cell signalling; insecticide;	
XX	OS	pharmaceutical; gene; ss.	
XX	PN	Drosophila melanogaster.	
XX	FD	WO200171042-A2.	
XX	27-SEP-2001.		

```

PF 23-MAR-2001: 2001WO-US09231.
XX
XX 23-MAR-2000: 2000US-191637P.
PR 11-JUL-2000: 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PMD, Myers EW;
XX
XX MPI: 2001-656860/75.
DR P-PSDB: ABB70184.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1: SEQ ID NO 37343: 21bp + Sequence Listing: English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB16351), expressed DNA
CC sequences (AB161737-AB162072),
CC (AB161737-AB16175) and the encoded proteins
CC
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 1300 BP; 346 A; 313 C; 369 G; 272 T; 0 other;
SQ
Query Match 20.0%; Score 276; DB 23; Length 1300;
Best Local Similarity 55.2%; Pred. No. 8, 7e-70;
Matches 581; Conservative 0; Mismatches 465; Indels 6; Gaps 2;
QY 37 ATGGGCATCAAGGTTTGACGAACCTGCTGGCGACAATGCCGCCAAGCGCATGAAGAG 96
DB 1 ATGGGAATTTTGGCTATCTAAGCTCATTTGCGGCTGCGCCACAGGCAATTCGCGAA 60
QY 97 CAGAAGTTCGAGACTACTTCGGCCGCAAAATCGCCGTCAGCGCCAGCATATATAC 156
DB 61 AGTGAGATGAAGCATTTTTCGTCGCAAGTAGCAATGATGCTATGCTGTAC 120
QY 157 CAGTTCCTGATGTTGTAAGACAGCATGGAATCTCACAATATACCTGTGA 216
DB 121 CAGTTCCTGATGCGCGGCGCTCCGAAGGC--GCCAGTTGGCCACGTAATGTGAT 177
QY 217 GTCACGTACTATTTGCAAGGAATGTTCAACCGGACAATTAATTAAGCGGGAATC 276
DB 178 CCCACGTCCCACTTAATGGCATGTTCTACCGCACCATCGATGCTGACAAAGGAATC 237
QY 277 AAGCGAGTTTATGTTTGTGTAAGCAAGCCCTGATGTAAGAAACAAGTTGCTTAA 336
DB 238 AAGCCGGATATGTCCTTCGACGGAAGCCACAGATCTCAAGTCCGCTGAGCTGAGAG 297
QY 337 AGATFACCTAAAAGAGATGATGCAACCAAGATCTGACTAGAGGACATAGAGTAGAGAT 396
DB 298 CGCGCGCGCGCGCGGAGGAAGCAAGCACTGAAGCGCGCCGCCCGGCAAGAT 357
QY 397 AAGATGCGATTTGAAATATGAGCAAGAGAGCTAAAGGTCACAGGCAACAACGAA 456
DB 358 GATGCGGATATGAAAGTTTAATCCCGCATTTGTCGCGGTAAAGAGAGACGCCAAA 417
QY 457 GATTGTAAAGCGCTATTAAACATTATGGGGTCTCTGTTGTAGAGCACCTTCTAGACA 516
DB 418 GAGGCGAAGAACTGCTACACTTAATGGGTGCTCTTATGTAGTACACCGGCAAGCG 477
QY 517 GAAGCAGATGTGACGCCCTTTGCAATAAGCATAGGTGCTGTTCTCAGAAAGT 576
DB 478 GAGGCGCGAGTGTGCGCTGTGTGAAAGCTGAAGGTTATGCCACGCCGAGGAGAT 537
QY 577 ATGAGCTCCCTACTTTTGGGGCTCCACGSGTCTCTTGTCATTTAATGATCCAAGTTC 636

```

```

DB 538 ATGATGCCCTTACATTTGGATCTACAAACTGTGAGATACCTTACCTACAGCGGCA 597
QY 637 AAGAAATACCTGTGTAAGAAATTTGANGTGGCCAAAGTTTGGAGAAGCTTGAACAC 696
DB 598 CGAAGATGCCCTCAAGAGCTTACAGTACGACAAAGCTTTGGAAGGTGCGCCATTAA 657
QY 697 ATGSAACAGTTGATTTGATTTGGATCTGCTGTGGATGTGATGATTTGATGATCAAA 756
DB 658 AATCGAGCTTATTTGATCTATGTTCTGCTGGTTGCGATTACTGTAGACATCAAG 717
QY 757 GGTATCGGGGCGCAACAGCTCTGAACCTTATGTCACATGGGTCCATGAAAGCATC 816
DB 718 GGTATGACCCCAACAGCGCATGCAACTGATCAACACCTATGCGATATAGACTAAT 777
QY 817 TTGGAATCTTAATAAGACGATATCAATTCCTGAGGACTGGCTTACCAAGAAGCT 876
DB 778 CTGATTAACCTTGACTTACGAAATACACCGTGGCCGAGAACTGGAAGGTGGC 837
QY 877 CGACGCTTTCAGAGACCTTAATGTAC--ATTGATATTCCTGAGCTAAATGACT 933
DB 838 CGGGAACCTTTCATGCAACCGAGGTACCTGATGCGCACTCCATAGATTCAAATGGGTC 897
QY 934 GCACCTGATGAGAGGCTCATATAGTTTCTGGTAAAGTATATGTTTCAACGAAGAT 993
DB 898 GACCGGATGAGAGAGGCTTGTCAAGTTTCTGCGCGGACCGGCACTTCAACGAAG 957
QY 994 CGSGTGCAAGGCCATAGAGATGATCAATTCGCCAAGTAATATGTCGAAGAGA 1053
DB 958 CGGTTTCGCAACGGTCCCAAAAGCTGATGAATTCAGAGGCCGACCTCAGGTGAGA 1017
QY 1054 CTCGAGTCTTTTTCAGAGCAACTGCCACAC 1085
DB 1018 CTCGATAGCTTCTTTAAGCACTGCCGAC 1049
RESULT 9
AAx02108
ID AAx02108 standard; cDNA; 1930 BP.
AC AAx02108;
XX
XX 23-APR-1999 (first entry)
DT
XX
DE Mouse FEN-1 cDNA.
XX
XX FEN-1; mouse; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage; ss.
XX
XX Mus sp.
XX
XX US874283-A.
XX
XX PD 23-FEB-1999.
XX
XX PF 30-MAY-1995; 95US-0455968.
XX
XX PR 30-MAY-1995; 95US-0455968.
XX
XX (HARR/) HARRINGTON J J.
PA (SITE/) HSTEH C.
PA (LIEB/) LIEBER M R.
XX
XX Harrington JJ, Hsieh C, Lieber MR;
XX
XX MPI: 1999-179985/15.
DR P-PSDB: AAW92505.
XX
XX DNA encoding flap endonuclease polypeptides - useful for producing
PT e.g. recombinant polypeptides
PS Disclosure; Fig 2B; 58pp; English.
XX

```

CC This sequence encodes a mouse FEN-1 (flap endonuclease) protein. This  
 CC protein can be used in methods for detecting a pathological condition in  
 CC a patient, for diagnostic purposes, for screening for antineoplastic  
 CC agents and carcinogens, for diagnostic staging of neoplasia, for  
 CC producing recombinant flap endonuclease for use as research or  
 CC diagnostic reagents, for producing antibodies reactive with the novel  
 CC polypeptides, for producing transgenic nonhuman animals expressing the  
 CC novel polypeptides encoded by a transgene. The invention also provides  
 CC novel molecular cloning techniques and reagents involving cleavage of  
 CC a flap or nick with a flap endonuclease.

XX Sequence 1930 BP; 482 A; 459 C; 553 G; 436 T; 0 other;

Query Match 19.7% Score 272.2; DB 20; Length 1930;

Best Local Similarity 55.1%; Pred. No. 1.4e-68; Mismatches 488; Indels 24; Gaps 4;

Matches 629; Conservative 0; Mismatches 488; Indels 24; Gaps 4;

QY 37 ATGGGCAATCAAGGCTTGAAGCAAGTCTGCGGCAAAATCGCCCGAAGCGGATGAAGAG 96  
 DB 1 ATGGGCAATCAAGGCTTGAAGCAAGTCTGCGGCAAAATCGCCCGAAGCGGATGAAGAG 60  
 QY 97 CAGAGTTCAGAGTCTGCTGCGGCAAAATCGCCCGAAGCGGATGAAGAGATATAC 156  
 DB 61 AATGACATCAAGAGTCTGCTGCGGCAAAATCGCCCGAAGCGGATGAAGAGATATAC 120  
 QY 157 CAGTTCCTGATTTAGTGAAGCAAGGATGAAGTCTCACAATGAAGCTGGTGA 216  
 DB 121 CAGTTCCTGATTTAGTGAAGCAAGGATGAAGTCTCACAATGAAGCTGGTGA 177  
 QY 217 GTGCTGATTTGCAAGGAATGTTCACCGGCAATATATACAGAGGCGGAATC 276  
 DB 178 ACCACCGAGCTGATGGGCAATGTAT-----GGCAACCATCCGATGGAATGGCATC 231  
 QY 277 AAGCCATTATGTTTGTATGCAAGCTCTCTGATGAAGCAAGCAAGCTTGTAA 336  
 DB 232 AAGCTGTGATGCTTTGATGCAAGCAAGCTCTCTGATGAAGCAAGCAAGCTTGTAA 291  
 QY 337 AGATCTCAAAAGATATGCAAGCAAGTCTGCTGATGAAGCAAGTGAAGAT 396  
 DB 292 CCGAGTGAAGGCGCGCGGCGGATGAAGCAAGTCTGCTGATGAAGCAAGTGAAGAT 351  
 QY 397 AAGATGCGGATTAATAATGAGCAAGGATGAAGGTCACAGGCAAGCAAGCA 456  
 DB 352 GAGGAGGAGTGAAGATTCACCAAGAGCTGCTGAAGGTCACCAAGCAAGCAAGAT 411  
 QY 457 GATTGTAAGGCTATTAGACTTATGGGCTTCTGTTGAGAGCAAGCTTGTGAAGA 516  
 DB 412 GAGTGAAGGCTATTAGACTTATGGGCTTCTGTTGAGAGCAAGCTTGTGAAGA 471  
 QY 517 GAAGCAAAATGTCAGAGCTTGTGCAATAAGATTAAGCTGCTTTCAGAGAT 576  
 DB 472 GAGGCAAGCTGCTGCTGCGGCAAGGCTGCAAAAGTCTATGCTGCGGCAAGGAG 531  
 QY 577 ATGAGCTCCCTTCTTGTGGGCTCCAGCGTCTCTTCATTTAATGATCAAGTTCC 636  
 DB 532 ATGAGCTCCCTTCTTGTGGGCTCCAGCGTCTCTTCATTTAATGATCAAGTTCC 591  
 QY 637 AAGAAATTAAGTGAAGATTTGATGTCAGCAAGGTTTGAAGAGCTTGAAGCAAG 696  
 DB 592 AAGAAATTAAGTGAAGATTTGATGTCAGCAAGGTTTGAAGAGCTTGAAGCAAG 651  
 QY 697 ATGAGCAAGTTCATTTGATGTCATCTGTGATGATGATGATGATGATGATGATG 756  
 DB 652 CAGAGAGAGTTTGTGATGTCATCTGTGATGATGATGATGATGATGATGATGATG 711  
 QY 757 GGTATCGGGGGGCAAGAGCTGGAAGCTTATTTGCAAACTGGTCCATGAAGAGATC 816  
 DB 712 GGTATCGGGGGGCAAGAGCTGGAAGCTTATTTGCAAACTGGTCCATGAAGAGATC 771  
 QY 817 TTGAGAGATTTTAAAGACAGATATCAAAATTCCTGAGAGCTGCTTACCAAGAGCT 876  
 DB 772 GTGAGGCGGTGAGCCCAAGCAAGTACCCGTTCCAGAGAGTGGCTCCACCAAGAGAG 831

QY 877 CGAGCTTGTTCAGAGGCTTAA---TGTCAATGATATTCCTGAGCTAAATGAGACT 933  
 DB 832 CAGAGCTTGTTCCTGAGGAGGAGAGTATGAGACCGAGCTGTGAGCTGAAGTGAAGC 891  
 QY 934 GCAGCTGATGAGGAGGCTCATAGTTTCTGTAAGATTAATGTTTCAACGAGAT 993  
 DB 892 GAGCCAAATGAAAGATGTTGTAATTTATGTTGTTGTAAGAGCTTTTGTGAAGAG 951  
 QY 994 CGGTCACAAAGGCTTAAAGAGATCAATCTGCCAAGATTAATGTCGCAAGAGAG 1053  
 DB 952 CGAATTCGAGTGGGTGCAAGCGGCTGAGTAAGAGCCCGGAGGAGACCGAGAGC 1011  
 QY 1054 CTGAGTCTTCTTTCAGCCCACTGCAACACATGAGCAGCGCTTAAACGAGAGACT 1113  
 DB 1012 CTGATGATTTCTTCAAGGAGAGAGCTCATCTCTCAGC-----TAAAGCGC 1059  
 QY 1114 TCGATTAAGCAAGCAAGGCTGCAAGCAAGAAAGCGTGGTGAAGAGAGAA 1173  
 DB 1060 AAGGAGCCAAAGACCAAGGCGCTGCTAAGAGAAAGCAAGAGCTGGGAGCGGAG 1119  
 QY 1174 T 1174  
 DB 1120 T 1120

RESULT 10  
 AAX02109  
 ID AAX02109 standard; cDNA; 1149 BP.  
 AC AAX02109;  
 XX 23-APR-1999 (first entry)  
 DE Yeast FEN-1 cDNA.  
 XX  
 KW FEN-1, yeast; flap endonuclease; detection; diagnosis; carcinogen;  
 KW neoplasia; antineoplastic agent; cleavage; ss.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 PN US5874283-A.  
 XX  
 PD 23-FEB-1999.  
 XX  
 PP 30-MAY-1995; 95US-0455968.  
 PR 30-MAY-1995; 95US-0455968.  
 XX  
 PA (HARR/) HARRINGTON J J.  
 PA (HSIEH/) HSIEH C.  
 PA (LIEBER/) LIEBER M R.  
 XX  
 PI Harrington JJ, Hsieh C, Lieber MR;  
 XX  
 DR WPI; 1999-179985/15.  
 DR P-PSDB; AAM92506.  
 XX  
 PT DNA encoding flap endonuclease polypeptides - useful for producing  
 XX e.g. recombinant polypeptides  
 XX  
 PS Disclosure; Fig 3b; 58bp; English.  
 XX  
 CC This sequence encodes a yeast FEN-1 (flap endonuclease) protein. This  
 CC protein is used in a method to isolate novel human FEN-1 proteins for  
 CC detecting a pathological condition in a patient, for diagnostic purposes,  
 CC for screening for antineoplastic agents and carcinogens, for diagnostic  
 CC staging of neoplasia, for producing recombinant flap endonuclease for use  
 CC as research or diagnostic reagents, for producing antibodies reactive  
 CC with the novel polypeptides, for producing transgenic nonhuman animals  
 CC expressing the novel polypeptides encoded by a transgene. The invention  
 CC also provides novel molecular cloning techniques and reagents involving  
 CC cleavage of a flap or nick with a flap endonuclease.  
 XX



QY 298 GGCAAGCCTCCTGATATAGAAACAAGAGCTTGTAAAGATCTCAAAAGAGATGAT 357  
 Db 1321 GGAAAGCCACCAATCTCAAGTCCGGTGGAGCTGGCCAGCGCGGCGGAGGGAGGA 1380  
 QY 358 GCAACCAAGATCTGACTAGGCAAGTAGAGATTAAGATGCAATGTAATAATTTG 417  
 Db 1381 GCGGAGAGAGCACTGAAGGCGGACCGAGTGCAGGAGATGATCCGGTATCGAAAGTTT 1440  
 QY 418 AGCAAGAGAGCACTGAAGGCTCACAGAGCAACAGAGATGTAAAGGGTATTAGA 477  
 Db 1441 AATGCCGATTTGGTCCGGGTACCAAGAGACAGCCAAAGAGGCCAAGAACTGCTACA 1500  
 QY 478 CTATGGGGTCTCTGTTTGAAGGCACTTTGAAGCAGAAACAGATGTCCACCTT 537  
 Db 1501 CTAATGGGTGTCCTTATGTTGATGACCGCTGCAGAGCCGAGGCCAGTGTCCCTCTG 1560  
 QY 538 TGCATTAACAGATAGGTGTCGGCTGCTGAGAGATATGAGTACCTTTTGGG 597  
 Db 1561 GTGAAGCTGAAGAGTTTATGCCACCGCCAGGAGATATGATCCCTCACATTCGA 1620  
 QY 598 GCTCCACGCTTCTTCTGTCATTTAATGATCAAGTTCCAAAGAAATACCTGTGATGA 657  
 Db 1621 TCTACAAAGCTGTGATATCTTACTACAGCGAGGACAGAAAGATGCCCTCAAGGAG 1680  
 QY 658 TTTGATGTTGCAGGTTTGGAGAGCTTGAACCTGACATGACATGACATTTGATTGG 717  
 Db 1681 TTGAGCTACGACAGCTGTTGGAGGCTGGCCATTACATCGAGATTCATGATCTA 1740  
 QY 718 TGCATCTGTGTGATGATGACTATTTGATATGATCAAGGATTCGCGGGGCAACAGCT 777  
 Db 1741 TGTATTTCTGGGGTTCGAGTACTGAGAGCATCAAGGGTATTTGAGCCCAAGGAGCG 1800  
 QY 778 CTGAAGCTTATGCTCAACATGGTCCATATGAAGATCTCTGAGATGTTAATAAGC 837  
 Db 1801 ATCGAAGCTGATCAACACTATCGGATATAGAGACTATTTGATATCTGAGCTCTAGC 1860  
 QY 838 AGATATCAAAATCTCTGAGAGCTGSCCTTACCAAGAGCTGAGCTGTTTCAAGAGCT 897  
 Db 1861 AAATACACCGTGGCCGAGAACTGAAGTGGCGGGGAACTCTTCAATCGAACCG 1920  
 QY 898 AATGTCAC---ATGATATTTCCGAGCTTAATGAGTGCAGTCCGATGAGAGGGCTTC 954  
 Db 1921 GAGGTAGCTGATGCCGACTCCATAGATCTCAATGGGTGAGCTCGGATGAGAGGGCTT 1980  
 QY 955 ATAAATTTCTGTTAAAGATTAATGTTTCAACGAAAGATCGGTTGACAAAGGCCATAGAG 1014  
 Db 1981 GTCAAGTTTCTCTGCGGCGACCGGCAAGTTCACGAAAGAGCGGCTTGGCAACGGTGCAAA 2040  
 QY 1015 AAGATCAAAATCTGCCAGATTAATGTCGCAAGAGAAAGTGCAGTCTTTTCAAGGCA 1074  
 Db 2041 AAGCTGATGAATCCAAAGCAGGCCCGACACTCAGGTGAGCTGATAGCTTTTAAGACA 2100  
 QY 1075 ACTGCCACAC 1085  
 Db 2101 CTGCCACGAC 2111

## RESULT 12

ABL10094 standard; cDNA: 5471 BP.

XX ABL10094  
 AC ABL10094:  
 XX 26-MAR-2002 (first entry)  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 24764.  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.

XX 27-SEP-2001.  
 PD 23-MAR-2001; 2001WO-US09231.  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 PA (PEKE ) PE CORP NY.  
 XX Venter JC, Adams M, Li PWD, Myers EM;  
 PI WPI; 2001-656860/75.  
 DR P-PSDB; ABB65991.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 PS Claim 1; SEQ ID NO 24764; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB101845), expressed DNA  
 CC sequences (AB101840-AB101845) and the encoded proteins  
 CC (AB101840-AB101845).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 5471 BP; 1422 A; 1223 C; 1325 G; 1501 T; 0 other:  
 Query Match 12.5%; Score 173.2; DB 23; Length 5471;  
 Best Local Similarity 54.0%; Pred. No. 1.6e-39;  
 Matches 377; Conservative 0; Mismatches 318; Indels 3; Gaps 1;  
 QY 391 GGAGATTAAGATCGATTTGAAAATTTGACACAGAGACTGTAAAGTCAACAGCAAC 450  
 Db 4 GGAGATATCCCGCTATCGAAAAGTTATCCGATTTGTCGGGTACCAAGAGCAGC 63  
 QY 451 AACGATTTGTAACGCTATTAGACTTATGAGGGTCTGTTGTAGAGCACTCTT 510  
 Db 64 GCCAAGAGGCCCAAGAACCTGCTACACATATGAGGCTGCTATGATGATGACCCGTC 123  
 QY 511 GAACAGACAGCAATGTGAGAGCCCTTTGCATTAACGATTAAGTGTGCTGTGCTCA 570  
 Db 124 GACGAGAGGCCAGTGTGCTGCTGTGTAAGCTGGAAGGTTTATGCCACCCGACG 183  
 QY 571 GAAGATATGACTCCCTTACTTTTGGGCTCCACGCTTCCATTTATGATGATCA 630  
 Db 184 GAGATATGATGATCCCTTACATATGATGATGATGATGATGATGATGATGATGATG 243  
 QY 631 AGTTCCAGAAATTAACCTGTGATGATGATGATGATGATGATGATGATGATGATG 690  
 Db 244 GAGGACAGAAAGATGCGCCGCAAGAGATGACATGACAGAGCTTTGGAAGTGTGCC 303  
 QY 691 CTCACATGAGCAAGTTCAATTTGATGATGATGATGATGATGATGATGATGATGATG 750  
 Db 304 ATTAACATGATGAGAGTTCAATTTGATGATGATGATGATGATGATGATGATGATG 363  
 QY 751 ATCAAGGATATGAGGGGSCCAACAGCTGAGAACTTATTCGCAACATGAGGATGAG 810  
 Db 364 ATCAAGGATATGAGAGCCCAAGAGATGAGAACTGATGATGATGATGATGATGATG 423  
 QY 811 AGCATCTTGAGATCTTATATAAGACAGATATCAAAATCTGAGAGCTGCTTACCAA 870  
 Db 424 ACTATTTCTGATTAACCTGAGCTCTAGCAAAATACACGCTGCGAGAACTGAGAACT 483  
 QY 871 GAAGCTGAGAGCTTTGTTCAAGAGGCTATGTCATGATGATGATGATGATGATGATG 927



Db 484 GTGGCGGGAACTTTCATCGAACGGAGTAGCTGATGCCAGCTCATGATCTCAAA 543  
 QY 928 TTGACATGCACCCATGATGAGAGGGTCCATCAATTCCTCGTGAAGATAATAGTTTCAC 987  
 Db 544 TGGGTGAGCGCGATGAGAGGGCTTGTCAATGTTCTGTGGCGCCGACGCTTCAAC 603  
 QY 988 GAAGATCGGGTGCACAGGCGCATAGAGAAGATCAATCTGCCAAGATTAATTCGTGCAA 1047  
 Db 604 GAAGAGCGGCTTCGCAACGCTGCAAAAAAGCTGATGAATCCCAAGCCAGACTCTAG 663  
 QY 1048 GGAAGACTGAGTCTTTCATGAGCCCACTGCCACCAC 1085  
 Db 664 GTGAGCTGATGCTTCTTTAAGACACTGCCACAC 701

RESULT 13  
 AAT76685  
 ID AAT76685 standard; DNA; 1023 BP.  
 AC AAT76685.

DT 14-APR-1998 (first entry)  
 DE Pyrococcus furiosus FEN-1 endonuclease gene coding sequence.  
 XX

KM Nucleic acid cleavage; 5' nuclease; DNA cleavage; RNA cleavage;  
 KM Invader directed cleavage; FEN-1; endonuclease; ds.

XX Pyrococcus furiosus.

PN MO9727214-A1.

PD 31-JUL-1997.

XX 22-JAN-1997; 97MO-US01072.

PR 02-DEC-1996; 96US-0759038.

PR 24-JAN-1996; 96US-0599491.

PR 12-JUL-1996; 96US-0682853.

PR 29-NOV-1996; 96US-0756386.

PR 02-DEC-1996; 96US-0758314.

XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

PA Brow MAD, Dahlberg JE, Hall JG, Kaiser KM, Lyamichev VI;  
 PI Olive DM, Prudent JR;

PT Thermostable structure-specific nuclease(s) - used for detection and  
 PT characterisation of nucleic acid sequences and variations in nucleic  
 PT acid sequences

XX Example 28; Page 283-285; 457bp; English.

XX This sequence comprises the coding region of the gene encoding

CC Pyrococcus furiosus (Pfu) FEN-1 endonuclease (see AAM24216).

CC It was obtained by PCR amplification (see AAT76682-83). Large

CC scale preparation of recombinant Pfu FEN-1 was performed using

CC E. coli as host. Pfu FEN-1 is a thermostable enzyme. It can be

CC used in novel methods for the detection and characterisation of

CC nucleic acid sequences and variations in nucleic acid sequences.

XX Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other;

Query Match 11.6%; Score 159.6; DB 18; Length 1023;  
 Best Local Similarity 54.6%; Pred. No. 5.9e-36;  
 Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 86 CGATGAGAGAGAGATTGAGAGCTTCTGCGCGCAAAATCCCGTCGACGCGACAGA 145  
 Db 29 CAAGAAAGAAATTTAGTAGAAACCTATACGGGAAAAAATCCATTCGACGCTCTTA 88

QY 146 TGAGCATATACCGATTCTCTGATTTGATTGGAAGCAGCGCATGGAACCTTCACAAATG 205  
 Db 89 ATGCATATCTACCAATTTTGTCCACATATAGCAGAAAGATGGAATCTTATGATTT 148  
 QY 206 AAGCTGTGCAAGTCACTAGTATTTGCAAGGAATGTTCAACCGGCAATAGATTACTCG 265  
 Db 149 CAAAGGGTGAATTAACCTCCACCTTAAGCGGGCTCTTTACAGAGCAATTAACCTAATG 208  
 QY 266 AAGCGGAAATCAAGCCAGTATGTTTGTGATGGCAAGCCCTCGTATATGAAGAAACAG 325  
 Db 209 AGCTGGAAATTAACCTGATGTTTGTGATGGAAGCTCCAGATTAATGAAGAAAG 268  
 QY 326 AGCTGTAAAGATTAAGTCAAAAGAGATGCAACCAAGATCTGACTGAGGCAATG 385  
 Db 269 AGCTGCAAAAG 328  
 QY 386 AGGTGAGAGATTAAGATGCAATTTGAAATTTGACCAAGAGAGAGAGAGAGAGAGAG 445  
 Db 329 AAAAGAGAGATTAAGATGCAATTTGAAATTTGACCAAGAGAGAGAGAGAGAGAGAG 388  
 QY 446 AACACAGCAGAGATTGTAAGGCTATTAGACTTATGAGGGGTTCTGTTGAGAGGAC 505  
 Db 389 TCGTCAATCGAGATGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448  
 QY 506 CTTTGAAGCAG 565  
 Db 449 CTAGCGAGGAG 508  
 QY 566 CTTGAGAGATTAAGTCAATTTGAGGCTTCTGAGGCTGAGAGAGAGAGAGAGAGAG 625  
 Db 509 CTAGTCAATGATTAAGTCAATTTGAGGCTTCTGAGGCTGAGAGAGAGAGAGAGAGAG 568  
 QY 626 ATCCAGTTCAGAGAAATTAAGTCAATTTGAGGCTTCTGAGGCTGAGAGAGAGAGAG 667  
 Db 569 TAACAG 610

RESULT 14  
 AAV65840  
 ID AAV65840 standard; DNA; 1023 BP.  
 XX

AC AAV65840;

DT 02-FEB-1999 (first entry)

DE Pyrococcus furiosus FEN-1 endonuclease gene ORF.

KM Nucleic acid detection; multiple sequential invasive cleavage;  
 KM FEN-1; endonuclease; nuclease; ds.

XX Pyrococcus furiosus.

XX WO9842873-A1.

PD 01-OCT-1998.

PF 24-MAR-1998; 98MO-US05809.

PR 24-MAR-1997; 97US-0823516.

PA (THIR-) THIRD WAVE TECHNOLOGIES INC.

PI Brow MAD, Hall JG, Kwiatkowski RW, Lyamichev VI;  
 PI Mast AL, Vavra SH;

PT WPI; 1998-557036/47.

DR P-PSDB; AAM79970.

XX Detecting target nucleic acid by sequence-specific cleavage of  
 PT complex with two specific oligonucleotides - used to detect  
 PT cytomegaloherpes DNA

XX

PS Example 28b: Page 316-317; 524pp; English.

XX This is the nucleotide sequence of the open reading frame encoding  
CC the FEN-1 endonuclease (see AAW/970) of *Pyrococcus furiosus* (Pfu).  
CC It was obtained from genomic DNA by PCR amplification (see  
CC also AAW6838-39). The PCR product has been ligated into vector  
CC pRC99a, and FEN-1 was expressed in *E. coli* cells. The invention  
CC relates to means for the detection and characterisation of nucleic  
CC acid sequences, and variations in nucleic acid sequences. It also  
CC relates to methods for forming a nucleic acid cleavage structure on  
CC a target sequence and cleaving this structure in a site-specific  
CC manner, preferably using a thermostable structure-specific nuclease  
CC such as FEN-1. Cleavage of the cleavage structure by the  
CC nuclease indicates the presence of specific nucleic acid sequences  
CC or specific variants. The invention further relates to methods for  
CC the separation of nucleic acid molecules based on charge, methods  
CC for the detection of non-target cleavage products via the formation  
CC of a complete and activated protein binding region, and methods for  
CC the detection of nucleic acid from various viruses (e.g. human  
CC cytomegalovirus) in a sample. The method amplifies the detection  
CC molecule rather than the target itself, is less subject to  
CC contamination than exponential amplification processes, and allows  
CC many targets to be analysed in a single reaction.

XX Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other;

SQ

Query Match 11.6%; Score 159.6; DB 19; Length 1023;  
Best Local Similarity 54.6%; Pred. No. 5.9e-36;  
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 86 CGATGAAGGAGCAGACGAGCTGCTGCGCCGCAAAATCGCGTGCAGCCAGCA 145  
DB 29 CAAAGGATGAGATATGCTTCCCACTTAAGCGGAGGCTTTTACAGACATTAACCTATG 88  
QY 146 TGACCATATACCATCTCCGATTTGATGAGCAAGCAGCATGAACTCTCACAATG 205  
DB 89 ATGCATATCTACCAATTTTGTCCCAATTAAGACAGAAATGAGTACCTCTATGATTT 148  
QY 206 AAGCTGAGGAGTACATTCATTTGCAAGGATGTTCAACCGGACCAATTAAGTTACTGG 265  
DB 149 CAAAGGATGAGATATGCTTCCCACTTAAGCGGAGGCTTTTACAGACATTAACCTATG 208  
QY 266 AAGCGGATGAGCAGCTTTATTTTATGAGCAAGCCTCTGATGAGCAAGCAAG 325  
DB 209 AGGCTGAGATTAACCTCTGATTTTATGAGCAAGCCTCTGATGAGCAAGCAAG 268  
QY 326 AGCTTCAAAAGATTAACCAAAAGAGATGATGCAACCAAGATGAGTGAAGAGAG 385  
DB 269 AGCTGAGAAAG 328  
QY 386 AGGTAGAGATTAAGATGCGATTTGAAATTTGAGCAAGAGAGAGAGAGAGAGAGAG 445  
DB 329 AAAAAAGGAGATTAAG 388  
QY 446 AACACAGAGAGATTTTAACGCTTATTAAGACTTATGAGGAGTTCGTTGAGAGAGC 505  
DB 389 TGCTCATGAGAGATGCAAAAAAATCTTAGAGCTTATGAGCAATTTCTATAGTTCAGAGC 448  
QY 506 CTTCGAAAGCAGAGAGAGATGTCAGCCCTTTGATTAAGATTAAGATTAAGATTAAG 565  
DB 449 CTAG 508  
QY 566 CTTCGAGATATGAG 625  
DB 509 CTAGCAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 568  
QY 626 ATCCAGATTCAGAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 667  
DB 569 TAACAG 610

RESULT 15  
AAV53951

ID AAV53951 standard; DNA; 1023 BP.

XX AC AAV53951;

XX DT 21-DEC-1998 (first entry)

XX DE Nucleotide sequence of the Pfu FEN-1 endonuclease.

XX KW Pfu FEN-1 endonuclease; Tag gene; structure-specific nuclease;  
KW mutant; DNA polymerase; bacteria; fungi; protozoa; RNA virus;  
KW hepatitis C virus; HCV; thermostable; ds.

XX OS *Pyrococcus furiosus*.

XX FH Key Location/Qualifiers  
FH CDS 1..1023  
FT /\*tag= a  
FT /product= "Pfu FEN-1 endonuclease"

XX PD 04-JUN-1998.

XX PE 26-NOV-1997; 97MO-0521783.

XX PR 02-DEC-1996; 96DS-0758314.  
PR 29-NOV-1996; 96DS-0757653.

XX PA (THIR-) THIRD WAVE TECHNOLOGIES INC.

XX PI Kaiser MW, Lyamichev VI, Lyamicheva N;  
PI WPI: 1998-322748/28.  
DR P-PSDB; AAW5940.

XX PT Thermostable structure-specific nuclease(s) derived from mutant DNA  
PT polymerase(s) - useful for detecting mutant allele(s) or strains of  
PT microorganisms

XX PS Example 29: Pages 278-279; 472pp; English.

XX CC This is the nucleotide sequence encoding the Pfu FEN-1 endonuclease,  
CC used in the method of the invention. In this process thermostable  
CC structure-specific nucleases are derived from mutant DNA polymerases,  
CC which can be used for detecting mutant alleles or strains of  
CC microorganisms. The structure-specific nucleases can be used in  
CC mixtures, compositions and kits to treat nucleic acid, e.g. for  
CC detection of wild type and mutant alleles of genes, for detection  
CC and/or identification of strains of microorganisms such as bacteria,  
CC fungi, protozoa, especially for detection of RNA viruses such as the  
CC hepatitis C virus (HCV).

SQ

Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other;

Query Match 11.6%; Score 159.6; DB 19; Length 1023;  
Best Local Similarity 54.6%; Pred. No. 5.9e-36;  
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 86 CGATGAAGGAGCAGACGAGCTGCTGCGCCGCAAAATCGCGTGCAGCCAGCA 145  
DB 29 CAAAGGATGAGATATGCTTCCCACTTAAGCGGAGGCTTTTACAGACATTAACCTATG 88  
QY 146 TGACCATATACCATCTCCGATTTGATGAGCAAGCAGCATGAACTCTCACAATG 205  
DB 89 ATGCATATCTACCAATTTTGTCCCAATTAAGACAGAAATGAGTACCTCTATGATTT 148  
QY 206 AAGCTGAGGAGTACATTCATTTGCAAGGATGTTCAACCGGACCAATTAAGTTACTGG 265  
DB 149 CAAAGGATGAGATATGCTTCCCACTTAAGCGGAGGCTTTTACAGACATTAACCTATG 208  
QY 266 AAGCGGATGAGCAGCTTTATTTTATGAGCAAGCCTCTGATGAGCAAGCAAG 325  
DB 209 AGGCTGAGATTAACCTCTGATTTTATGAGCAAGCCTCTGATGAGCAAGCAAG 268

Tue May 27 16:02:22 2003

us-09-805-311-5.std.rng

Page 15

326 AGCTCTCAAGATGATCTCAAAAAGAGATATGTCACCAAAATCTGACTAGGACATAG 385  
 326 AGCTCTCAAGATGATCTCAAAAAGAGATATGTCACCAAAATCTGACTAGGACATAG 385  
 269 AGCTCGAAAAAAGAGAGACACGAGAGAGACTCAACAAAAAGTGGAGAACACACTTG 328  
 386 AGGTAGAGAAATTAAGATCGATTGAAAAATTTGACCAAGAGAGACTGTAAGGTCACAAAGC 445  
 329 AAAAAAGAGATATAGAGAGAACCAAGAAAAATGCTCCCAAAAGACCAACGAGGTAAATGAAA 388  
 446 AACACACAGCAACATTTGTAACGGCTATTAAAGACTTTTGAGGGGCTCTCTTTAGAGGCAC 505  
 389 TGCTCATCGAGGATGCAAAAAAACCTTTAGAGCTTATGGGAATTCCTATAGTGTCAAGCAC 448  
 506 CTTCCTAAGCAGCAACACAGAAATGTGTCAGCCCTTTGCATTAACGATTAAGGTGTTCCCTGTG 555  
 449 CTAGCGGAGAGAGGCCCAAGCTGCTATATATGCGCCCAAAAGAGAGGAGCGTGTATGCTGCG 508  
 566 CTTCAGAGATATGAGCTCCCTACTTTTGGGGCTCCACAGGTTCCCTGCTATTAAATGCG 625  
 509 CTAGTCAAGATTACGATTCCTACTTTTGGAGCTCCAGAGACTGTGTAGAAACTTAACAA 568  
 626 ATCCAAAGTCCAAAGAAATTAAGCTGTGATGAGCAATTTGATGTTG 667  
 569 TTACAGAAAAAGAAAGTTGCTGCTGGAAAAAATGTCATACGTCG 610

```
Search completed: November 5, 2002, 11:01:08
Job time : 179.876 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 08:33:28 ; Search time 33.87 seconds  
(without alignments)  
10015.359 Million cell updates/sec

Title: US-09-805-311-5

Perfect score: 1381  
Sequence: 1 cgaccacgcgcgcgcgcac.....ttgaaaaaaaaaaaaa 1381

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues.

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents NA: \*

- 1: /cgn2.6/prodata/1/lna/5A.COMB.seq:\*
- 2: /cgn2.6/prodata/1/lna/5B.COMB.seq:\*
- 3: /cgn2.6/prodata/1/lna/6A.COMB.seq:\*
- 4: /cgn2.6/prodata/1/lna/6B.COMB.seq:\*
- 5: /cgn2.6/prodata/1/lna/PCFUS.COMB.seq:\*
- 6: /cgn2.6/prodata/1/lna/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1381	100.0	1381	4	US-09-426-557-5
2	1350.2	97.8	1463	4	US-09-426-557-1
3	1348.6	97.7	1541	4	US-09-426-557-3
4	1334.2	96.6	1478	4	US-09-426-557-7
5	302.6	21.9	1144	2	US-08-455-968E-2
6	302.6	21.9	1144	2	US-08-455-968E-28
7	276.6	20.0	2033	2	US-08-455-968E-9
8	272.2	19.7	1930	2	US-08-455-968E-4
9	270.2	19.6	1149	2	US-08-455-968E-6
10	158.6	11.6	1023	2	US-08-757-653-175
11	158.6	11.6	1023	2	US-08-823-516-78
12	159.6	11.6	1023	3	US-08-759-038-114
13	159.6	11.6	1023	3	US-08-758-314-114
14	135.6	9.8	1032	4	US-09-146-319-1
15	135.6	9.8	1032	4	US-09-175-973-1
16	89.2	6.5	981	2	US-08-757-653-171
17	89.2	6.5	981	2	US-08-823-516-74
18	89.2	6.5	981	3	US-08-759-038-110
19	89.2	6.5	981	3	US-08-758-314-110
20	65.2	4.7	7218	1	US-08-232-463-14
21	47	3.4	455	1	US-08-636-928-4
22	37.6	2.7	289	4	US-09-007-005-17
23	37.6	2.7	289	4	US-09-244-796-17
24	37.6	2.7	2339	4	US-09-268-140-11
25	37.6	2.7	2505	4	US-09-268-140-11
26	37.6	2.7	2517	4	US-09-268-140-7
27	36.4	2.6	1412	1	US-08-097-831-1

28	34.4	2.5	3825	4	US-09-208-742-3	Sequence 3, Appl1
29	34.2	2.5	5361	4	US-08-973-462-2	Sequence 2, Appl1
30	34.2	2.5	6152	4	US-08-973-462-1	Sequence 1, Appl1
31	34	2.5	677	3	US-08-881-094-15	Sequence 15, Appl1
32	34	2.5	36741	4	US-09-301-665-3	Sequence 3, Appl1
33	33.8	2.4	3791	4	US-09-377-310-1	Sequence 1, Appl1
34	33.8	2.4	43804	4	US-09-171-461-1	Sequence 1, Appl1
35	33.6	2.4	1272	2	US-08-972-258-1	Sequence 1, Appl1
36	33.6	2.4	1272	4	US-09-263-128-1	Sequence 1, Appl1
37	33.6	2.4	2346	4	US-09-193-503B-4	Sequence 4, Appl1
38	33.6	2.4	2346	4	US-09-193-503B-5	Sequence 5, Appl1
39	33.6	2.4	2346	4	US-09-193-503B-8	Sequence 8, Appl1
40	33.6	2.4	72604	4	US-09-268-992-7	Sequence 7, Appl1
41	33.4	2.4	984	4	US-08-927-215-52	Sequence 52, Appl1
42	33.4	2.4	1858	3	US-08-742-185-96	Sequence 96, Appl1
43	33.4	2.4	10014	4	US-08-927-219-130	Sequence 130, Appl1
44	33.2	2.4	48974	4	US-08-920-422-17	Sequence 17, Appl1
45	33	2.4	23673	4	US-09-773-816-1	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-426-557-5  
; Sequence 5, Application US/09426557  
; Patent No. 6232527  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
; FILE REFERENCE: 0961  
; CURRENT APPLICATION NUMBER: US/09/426,557  
; EARLIER FILING DATE: 1999-10-22  
; EARLIER APPLICATION NUMBER: 60/112,332  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 1381  
; TYPE: DNA  
; ORGANISM: Zea mays  
; NAME/KEY: CDS  
; LOCATION: (37)...(1173)  
US-09-426-557-5

Query Match 100.0%; Score 1381; DB 4; Length 1381;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGACCCAGCGCGTCCGGCCACAGCCGCCGACGAGATGGGCAATCAAGGATTGACGANA	60
DB	1	CGACCCAGCGCGTCCGGCCACAGCCGCCGACGAGATGGGCAATCAAGGATTGACGANA	60
QY	61	CTGCTGCGCGGACATGCCGCCAAGCGCATGAAGAGCAGAGTTCCAGAGCTACTTCGGC	120
DB	61	CTGCTGCGCGGACATGCCGCCAAGCGCATGAAGAGCAGAGTTCCAGAGCTACTTCGGC	120
QY	121	CGCAAAATGCGCGTCCGGCCACAGCCGCCGACGATGATATACAGTTCTGATTGATGAAG	180
DB	121	CGCAAAATGCGCGTCCGGCCACAGCCGCCGACGATGATATACAGTTCTGATTGATGAAG	180
QY	181	ACAGGATGGAATCTCTCAATATGAAGCTGTGAATCTACTGATTTGCAAGGAATG	240
DB	181	ACAGGATGGAATCTCTCAATATGAAGCTGTGAATCTACTGATTTGCAAGGAATG	240
QY	241	TTCAACCGGACATTAAGATTACTGGAACCGGGAATCAAGCCAGTTATGTTTGTATGTC	300
DB	241	TTCAACCGGACATTAAGATTACTGGAACCGGGAATCAAGCCAGTTATGTTTGTATGTC	300
QY	301	AAGCCCTGATATGAAGAAACAGAGCTTGCTAAAGATATCAAAAAGAGATGATGA	360
DB	301	AAGCCCTGATATGAAGAAACAGAGCTTGCTAAAGATATCAAAAAGAGATGATGA	360

Db 301 AAGCTCTGATATGAGAAACAGAGCTGCTGCTAAAGTACTCAAAAAAGATGATGCA 360  
QY 361 ACCAAATGATGTGACGTAGAGAGTAGAGATTAAGATGGCATTGAAAAATTGAGC 420  
Db 361 ACCAAATGATGTGACGTAGAGAGTAGAGATTAAGATGGCATTGAAAAATTGAGC 420  
QY 421 AAGAGAGCTTAAGGTGACAAAGCAGACAGAGATTGTAAAGCGCTATTAAAGCTT 480  
Db 421 AAGAGAGCTTAAGGTGACAAAGCAGACAGAGATTGTAAAGCGCTATTAAAGCTT 480  
QY 481 ATGGGGTTCCTGTGTAGAGCACCCTTGTGAAGCAGAGAGATGTGCAAGCCCTTTGC 540  
Db 481 ATGGGGTTCCTGTGTAGAGCACCCTTGTGAAGCAGAGAGATGTGCAAGCCCTTTGC 540  
QY 541 ATTAACATAGAGTTCCTGCTGTCTGACAGAGATGAGCCTCCCTACTTTTGGGGCT 600  
Db 541 ATTAACATAGAGTTCCTGCTGTCTGACAGAGATGAGCCTCCCTACTTTTGGGGCT 600  
QY 601 CCACGGTTCCTGCTATTTAATGATCCAAAGTTCAGAAAAAATACCTGTGATGSAATT 660  
Db 601 CCACGGTTCCTGCTATTTAATGATCCAAAGTTCAGAAAAAATACCTGTGATGSAATT 660  
QY 661 GATGTTCCCAAGGTTTGGAGAGGCTTGAACTCAACCATGAGCCAGTTCAATTGATTGCG 720  
Db 661 GATGTTCCCAAGGTTTGGAGAGGCTTGAACTCAACCATGAGCCAGTTCAATTGATTGCG 720  
QY 721 ATCTGTGTGATGTGACTATTTGATGATCAATCAAAAGTATCGGGGGGCAAGCCTCTG 780  
Db 721 ATCTGTGTGATGTGACTATTTGATGATCAATCAAAAGTATCGGGGGGCAAGCCTCTG 780  
QY 781 AAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
Db 781 AAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 841 TATTAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db 841 TATTAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
QY 901 GTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Db 901 GTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
QY 961 TTCTGTGATTAAGATTAAGTGTTCACAGAGAGTGGGTGACAAAGGCTCATAGAGATG 1020  
Db 961 TTCTGTGATTAAGATTAAGTGTTCACAGAGAGTGGGTGACAAAGGCTCATAGAGATG 1020  
QY 1021 AAATCTGCCAAGATTAATGCTGCGCAAGAGACTGCAAGTCTCTTTTCAAGCCAACTGCC 1080  
Db 1021 AAATCTGCCAAGATTAATGCTGCGCAAGAGACTGCAAGTCTCTTTTCAAGCCAACTGCC 1080  
QY 1081 ACCACATCAGACCGCTAAAGAGAGAGAGTGGATTAAGAGAGAGAGAGAGAGAGAGAG 1140  
Db 1081 ACCACATCAGACCGCTAAAGAGAGAGAGTGGATTAAGAGAGAGAGAGAGAGAGAGAG 1140  
QY 1141 AACAG 1200  
Db 1141 AACAG 1200  
QY 1201 CGACTAG 1260  
Db 1201 CGACTAG 1260  
QY 1261 ACTGAG 1320  
Db 1261 ACTGAG 1320  
QY 1321 AAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
Db 1321 AAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
QY 1381 A 1381  
Db 1381 A 1381

RESULT 2  
US-09-426-557-1  
; Sequence 1, Application US/09426557  
; Patient No. 6232527  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
; TITLE OF INVENTION: thereof  
; FILE REFERENCE: 0961  
; CURRENT APPLICATION NUMBER: US/09/426,557  
; CURRENT FILING DATE: 1999-10-22  
; EARLIER APPLICATION NUMBER: 60/112,332  
; EARLIER FILING DATE: 1998-12-15  
; NUMBER OF SEQ. ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ. ID NO 1  
; LENGTH: 1463  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (85) ... (1221)  
US-09-426-557-1  
Query Match 97.8%, Score 1350.2, DB 4, Length 1463;  
Best Local Similarity 99.4%, Pred. No. 0;  
Matches 1355; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CGACCCAGCAGTCCGCGCCACAGAGCCGCGCAGAGAGATGGGATCAAGGGTTTGACGAA 60  
Db 49 CGCGCGCGCGCCACCGCCACAGAGCCGCGCAGAGAGATGGGATCAAGGGTTTGACGAA 108  
QY 61 CTCTGCGCGGACAAATGCGCGCCCAAGGCGATGAAGAGAGAGAGAGAGAGAGAGAGAG 120  
Db 109 CTCTGCGCGGACAAATGCGCGCCCAAGGCGATGAAGAGAGAGAGAGAGAGAGAGAGAG 168  
QY 121 CGCAAAATGCGCGTGGACCGCCAGATGAGATGATGATGATGATGATGATGATGATGAT 180  
Db 169 CGCAAAATGCGCGTGGACCGCCAGATGAGATGATGATGATGATGATGATGATGATGAT 228  
QY 181 ACAGGATGAG 240  
Db 229 ACAGGATGAG 288  
QY 241 TTCAACCGGACATTAAGATTAAGTGTTCACAGAGAGTGGGTGACAAAGGCTCATAGAG 300  
Db 289 TTCAACCGGACATTAAGATTAAGTGTTCACAGAGAGTGGGTGACAAAGGCTCATAGAG 348  
QY 301 AAGCCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
Db 349 AAGCCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 408  
QY 361 ACCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
Db 409 ACCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 468  
QY 421 AAG 480  
Db 469 AAG 528  
QY 481 ATGGGGTTCCTGTGTAGAGCACCCTTGTGAAGCAGAGAGATGTGCAAGCCCTTTGC 540  
Db 529 ATGGGGTTCCTGTGTAGAGCACCCTTGTGAAGCAGAGAGATGTGCAAGCCCTTTGC 588  
QY 541 ATTAACATAGAGTTCCTGCTGTCTGACAGAGATGAGCCTCCCTACTTTTGGGGCT 600  
Db 589 ATTAACATAGAGTTCCTGCTGTCTGACAGAGATGAGCCTCCCTACTTTTGGGGCT 648  
QY 601 CCACGGTTCCTGCTATTTAATGATCCAAAGTTCAGAAAAAATACCTGTGATGSAATT 660  
Db 649 CCACGGTTCCTGCTATTTAATGATCCAAAGTTCAGAAAAAATACCTGTGATGSAATT 708

```

RESULT 3
US-09-426-557-3
Sequence 3, Application US/09426557
Patent No. 6232527
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
FILE OF INVENTION: Thereof
FILE REFERENCE: 0961
CURRENT APPLICATION NUMBER: US/09/426,557
CURRENT FILING DATE: 1999-10-22
EARLIER APPLICATION NUMBER: 60/112,332
EARLIER FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 1541
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (79) ... (1215)
US-09-426-557-3

```

QY	1	CGACCCAGGCGTCGGCCACACAGCCCGCCGACGACGAGATGGGCATTCAGAGGTTTGGACGAA	60
Db	43	CGCGCGCCGCCACCCGGCCACACGCCGCCGACGACGAGATGGGCATTCAGAGGTTTGGACGAA	102
QY	61	CTGCTGGCGGACAAATGCGGCCCAAGGCGATGAAGAGCAAGATTGAGAGCTACTTGGCG	120
Db	103	CTGCTGGCGGACAAATGCGGCCCAAGGCGATGAAGAGCAAGATTGAGAGCTACTTGGCG	162
QY	121	CGCAAAATGGCGCTGACAGCCGACATGACATATPACAGTTTCCGATTTGATGGGAAG	180
Db	163	CGCAAAATGGCGCTGACAGCCGACATGACATATPACAGTTTCCGATTTGATGGGAAG	222
QY	181	ACAGGCATGAAACTCTCACAAATGAAAGCTGATGACACTAGTCAATTTGCAAGSATG	240
Db	223	ACAGGCATGAAACTCTCACAAATGAAAGCTGATGACACTAGTCAATTTGCAAGSATG	282
QY	241	TTCAACCCGCAAAATAGATTACTGGAAGGGGGAATCAAGCAGTTTATGTTTGTGATGC	300
Db	283	TTCAACCCGCAAAATAGATTACTGGAAGGGGGAATCAAGCAGTTTATGTTTGTGATGC	342
QY	301	AAAGCTCTGATATGCAAGAAACAAGACTGTCTAAAGATACTCAAAAAGAGATGATGA	360
Db	343	AAAGCTCTGATATGATGAAGAAACAAGACTGTCTAAAGATACTCAAAAAGAGATGATGA	402
QY	361	ACCAAAAGATCTAGACGAGCGCGTGAAGGTAGAGGATPAAAGATGGTGATGTAATAATGAGC	420
Db	403	ACCAAAAGATCTAGACGAGCGCGTGAAGGTAGAGGATPAAAGATGGTGATGTAATAATGAGC	462
QY	421	AAGAGCACTGTAAAGCTACAGAGCAACAACGAAATTTGAAAGCGCTATTAGACTT	480
Db	463	AAGAGCACTGTAAAGCTACAGAGCAACAACGAAATTTGAAAGCGCTATTAGACTT	522
QY	481	ATGGGGGTTCCGTGTGTGAGAGGACCTCTGCAAGCAAGCAGAAATGTGAGCCCTTTCG	540
Db	523	ATGGGGGTTCCGTGTGTGAGAGGACCTCTGCAAGCAAGCAGAAATGTGAGCCCTTTCG	582
QY	541	ATPAAAGATPAAAGGTGTGCTGTGCTTCCGATGAGATATGACCTCCCTACTTTGGGGCT	600
Db	583	ATPAAAGATPAAAGGTGTGCTGTGCTTCCGATGAGATATGACCTCCCTACTTTGGGGCT	642
QY	601	CGACGCTCCCTGCTGATTAATGAGATCCAGATGCCAAGAAATACCTGGATGAGATTT	660
Db	643	CGACGCTCCCTGCTGATTAATGAGATCCAGATGCCAAGAAATACCTGGATGAGATTT	702
QY	661	GATGTTGCCCAAGGTTTGGAGAGACCTTGACATCACATGAGCAAGTTTATGATTTGTGC	720
Db	703	GATGTTGCCCAAGGTTTGGAGAGACCTTGACATCACATGAGCAAGTTTATGATTTGTGC	762
QY	721	ATGCTGTGTGATGTGACTATTTGATAGATATCAAAAGTATGGGGGGCAACAAGCTCG	780
Db	763	ATGCTGTGTGATGTGACTATTTGATAGATATCAAAAGTATGGGGGGCAACAAGCTCG	822
QY	781	AAACTATTTCGCAACATGGGTCATAGAAAGCATCTTGGAGAAATCTTAATAAGACAGA	840
Db	823	AAACTATTTCGCAACATGGGTCATAGAAAGCATCTTGGAGAAATCTTAATAAGACAGA	882
QY	841	TATCAAAATTCCTGAGGACTGGCTTACCAAGAAAGCTCGAGCTTGTTCAGAGAGCTAAT	900
Db	883	TATCAAAATTCCTGAGGACTGGCTTACCAAGAAAGCTCGAGCTTGTTCAGAGAGCTAAT	942
QY	901	GTCACATTTGATATTCCTGAGACTAAATGGACATGCACTGATGAGAGAGGCTCATAGT	960
Db	943	GTCACATTTGATATTCCTGAGACTAAATGGACATGCACTGATGAGAGAGGCTCATAGT	1002
QY	961	TTCTCGTGAAGATATATGTTTCAACGAAAGATCGGCTGACAAAGCCATATGAGAAAGATC	1020
Db	1003	TTCTCGTGAAGATATATGTTTCAACGAAAGATCGGCTGACAAAGCCATATGAGAAAGATC	1062

```

1021 AAATCTGCCAAGATATAATCTGTGCAAGGAGACACTGAGTCTTTTTCAGGCCAAGCTGCC 1080
1063 AAATCTGCCAAGATATAATCTGTGCAAGGAGACACTGAGTCTTTTTCAGGCCAAGCTGCC 1122
1081 ACCACATCAGCACCCTGTAAACGGAAGGAGACTTGGATTAACAGCAAGGACGCTGGC 1140
1123 ACCACATCAGCACCCTGTAAACGGAAGGAGACTTGGATTAACAGCAAGGACGCTGGC 1182
1141 AACAGAAAAACAAGGCTGTGGAAGAAAGAAATATCTTGTGATGCTGGATGATCAACTA 1200
1183 AACAGAAAAACAAGGCTGTGGAAGAAAGAAATATCTTGTGATGCTGGATGATCAACTA 1242
1201 CGACTACGAAAGACGCGGTGGCGGTATCATCTGCTTAGATTAATTAACCTCCCTGTTTA 1260
1243 CGACTACGAAAGACGCGGTGGCGGTATCATCTGCTTAGATTAATTAACCTCCCTGTTTA 1302
1261 ACTCAGAGCTTTGGTAAAGTTGCGCCATGTTTCAAGGTGGGATGAGTATGTTGTTTG 1320
1303 ACTCAGAGCTTTGGTAAAGTTGCTCATGTTTCAAGGTGGGATGAGTATGTTGTTTG 1362
1321 AAGAGATTGGTGTACCAAGTAACAAAACCTTATCGCTGTTTTTT 1363
1363 AAGAGATTGGTGTACCAAGTAACAAAACCTTATCGCTGTTTTTT 1405

RESULT 4
US-09-426-557-7
: Sequence 7, Application US/09426557
: Patent No. 623527
: GENERAL INFORMATION:
: APPLICANT: Mahajan, Pramod B.
: TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
: FILE REFERENCE: 0961
: CURRENT APPLICATION NUMBER: US/09/426,557
: CURRENT FILING DATE: 1999-10-22
: EARLIER APPLICATION NUMBER: 60/112,332
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 7
: LENGTH: 1478
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (97)...(1233)
US-09-426-557-7

Query Match          96.6%: Score 1334.2; DB 4; Length 1478;
Best Local Similarity 98.7%: Pred. No. 0;
Matches 1345; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

1 CGACCCACGCGGTCCGGCCACAGCCGCCGACAGAGATGGGCATCAAGGTTTGACGAAA 60
61 CGCGCGCCGCCACCGCCACAGCCGCCGACAGAGATGGGCATCAAGGTTTGACGAAA 120
61 CTGCTGCGGACAAATGGCGCAAGGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
121 CTGCTGCGGACAAATGGCGCAAGGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 160
121 CGCAAAATCGCGCTGCGACGCGACATGAGCATCTACAGTTCCGATGATGGAAG 180
181 CGCAAAATCGCGCTGCGACGCGACATGAGCATCTACAGTTCCGATGATGGAAG 240
181 ACAGGCAATGGAACCTCTCAAAATGAGAGTGTGGAAGTCACTAGCATTTGCAAGATG 240
241 ACAGGCAATGGAACCTCTCAAAATGAGAGTGTGGAAGTCACTAGCATTTGCAAGATG 300
241 TTCAACGCGAACAATAAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGAATGC 300
301 TTCAACGCGAACAATAAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGAATGC 360
```

```

301 AAGCCTCTGATATGAAGAAACAAGAGCTTGTCTAAAGATACCTCAAAAAGATGATGCA 360
361 AAGCCTCTGATATGAAGAAACAAGAGACTTGTCTAAAGATACCTCAAAAAGATGATGCA 420
361 ACCAAGATCTGACTAGAGGAGTAGAGTAGAGATTAAGATTCGATTTGAAAAATTTAGC 420
421 ACCAAGATCTGACTAGAGGAGTAGAGTAGAGATTAAGATTCGATTTGAAAAATTTAGC 480
421 AAGAGACTGTAAAGGTCAACAAGGACACACAGCAAGATTTGAAACGACTTTAAGACT 480
481 AAGAGACTGTAAAGGTCAACAAGGACACACAGCAAGATTTGAAACGACTTTAAGACT 540
481 ATGGGGTTCCTGTTGTAGAGGACCTTGTGAAGCAAGATGTCAGCCCTTTGC 540
541 ATGGGGTTCCTGTTGTAGAGGACCTTGTGAAGCAAGATGTCAGCCCTTTGC 600
541 ATTAACGATTAAGGTGTCGCTGTGCTTGAAGATTAAGATGTCGCTTACTTTGGGGCT 600
601 ATTAACGATTAAGGTGTCGCTGTGCTTGAAGATTAAGATGTCGCTTACTTTGGGGCT 660
601 CCACGCTTCCTGCTGCTATTAATGATCCAGTTCGCAAGAAATACCTGTGATGAATTT 660
661 CCACGCTTCCTGCTGCTATTAATGATCCAGTTCGCAAGAAATACCTGTGATGAATTT 720
661 GATGTGCCAAGGTTTGGAGAGCTTGAACCTCACCATGAGACCACTTATGATTTGTC 720
721 GATGTGCCAAGGTTTGGAGAGCTTGAACCTCACCATGAGACCACTTATGATTTGTC 780
721 ATCTGTGTGATGTGACTATTTGTATGATCAATCAAAAGTATCGGGGGCAACAGCTCTG 780
781 ATCTGTGTGATGTGACTATTTGTATGATCAATCAAAAGTATCGGGGGCAACAGCTCTG 840
781 AAATTAATGTGTCAACAVGGGTCATAGAAAGCATCTTGGAAATCTTAATAAGACAGA 840
841 AAATTAATGTGTCAACAVGGGTCATAGAAAGCATCTTGGAAATCTTAATAAGACAGA 900
841 TATCAAAATTCCTGAGAGCTGCGCTTACCAAGAGGCTGCGCTGTTCAAGAGACCTAAT 900
901 TATCAAAATTCCTGAGAGCTGCGCTTACCAAGAGGCTGCGCTGTTCAAGAGACCTAAT 960
901 GTCAATTTGGATATTCCTGAGAGCTGCGCTTACCAAGAGGCTGCGCTGTTCAAGAGAT 960
961 GTCAATTTGGATATTCCTGAGAGCTGCGCTTACCAAGAGGCTGCGCTGTTCAAGAT 1020
961 TTCTGTGTAAGATTAATGTTTCAAGAGATCGGGTACAAAGGCCATAGAGAAATC 1020
1021 TTCTGTGTAAGATTAATGTTTCAAGAGATCGGGTACAAAGGCCATAGAGAAATC 1080
1021 AAATCTGCCAAGATTAATGTTTCAAGAGATCGGGTACAAAGGCCATAGAGAAATC 1080
1081 AAATCTGCCAAGATTAATGTTTCAAGAGATCGGGTACAAAGGCCATAGAGAAATC 1140
1081 ACCACATCAGCACCCTGTAAACGGAAGGAGACTTGGATTAACAGCAAGGACGCTGGC 1140
1141 ACCACATCAGCACCCTGTAAACGGAAGGAGACTTGGATTAACAGCAAGGACGCTGGC 1200
1141 AACAGAAAAACAAGGCTGTGGAAGAAAGAAATATCTTGTGATGCTGGATGATCAACTA 1200
1201 AACAGAAAAACAAGGCTGTGGAAGAAAGAAATATCTTGTGATGCTGGATGATCAACTA 1260
1201 CGACTACGAAAGACGCGGTGGCGGTATCATCTGCTTAGATTAATTAACCTCCCTGTTTA 1260
1261 CGACTACGAAAGACGCGGTGGCGGTATCATCTGCTTAGATTAATTAACCTCCCTGTTTA 1320
1261 ACTCAGAGCTTTGGTAAAGTTGCGCCATGTTTCAAGGTGGGATGAGTATGTTGTTTG 1320
1321 ACTCAGAGCTTTGGTAAAGTTGCGCCATGTTTCAAGGTGGGATGAGTATGTTGTTTG 1380
1321 AAGAGATTGGTGTACCAAGTAACAAAACCTTATCGCTGTTTTTT 1363
1381 AAGAGATTGGTGTACCAAGTAACAAAACCTTATCGCTGTTTTTT 1423
```



```

RESULT 5
US-08-455-968E-2
; Sequence 2, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-455-968E-2

Query Match      21.9%  Score 302.6;  DB 2;  Length 1144;
Best Local Similarity 56.6%;  Pred. No. 2.2e-80;
Matches 601;  Conservative 0;  Mismatches 454;  Indels 6;  Gaps 2;

```

```

QY 457 GATTGTAACGCGCTATTACACTTATGGGGTTCCTGTTGTAGAGCCACCTCTTGAAAGA 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 418 GAGTGCACAAACATCTGTGACCCATCATGGGCATCCCTTATCTTGATGCACCAAGTAGGA 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 GAACGAAATGTGCAGCGCTTTCATTAACGATAGGTGTGCTGTTCCTTCAGAAAGT 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 478 GAGGCCACGTGTCTGTGCTGCTGGTGAAGGCTGGCAAAAGTCTATGTCTGGCTACCGAGGAC 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 577 ATGAGTCCTTACTTTTGGGGCTCCAGCGTTCCTTCCTGATTTATAGATCCAAATTC 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 538 ATGACTGCTCTACCTTCGCGACGCCCTGTGTATGCAATGCAACCTGACCTCCAGTAGAGC 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 637 AAGAAATACCTGTGATGAAATTTGATTTGTGCGCAGGTTTGGAGAGCTTGAAGTCAAC 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 598 AAAAAGCTGCCAATCCAGAAATTCACACGAGCCGATCTCTCAGAGCTGGCGCTGAAAC 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 697 ATGACCAGTTCATTTGATTTGTGATCTCTGTGTGATGTGATTTGATGATGATCAAA 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 658 CAGAGACAGTTTGTGATCTGTGATCTGTGATCTGTGATGAGTACTGTGAGATTCGG 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 757 GGTATCGGGGGGCAAAAGAGCTGTGAACTTATTCGTCAACATGGGTCCATAGAAAGATC 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 718 GGTATTTGGGCCCAAGCGGCTGTGACCTCATTCAGAAACACAAAGCATCGAGAGATC 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 817 TTGAGAACTTTAATTAAGACATATCAAAATTCCTGAGAGATGGCTTACCAAGAGCT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 778 GTCCGGCGACTTGAACCCCAACAGTAATCCTGTGCGAAGAAATTTGCTCCACAGAGAGCT 837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 CGAGCGCTTGTCAAGAGAGCTTAATG---TCACATTGATATTCAGAGTAATAGTACT 933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 838 CACACAGCTCTTCTTGAACCTGAGGTGTGAGCCAGAGATCTGTGAGCTGAAGTGGAGC 897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 934 GCACCTGATGAGAGAGGTCTCATTAATTCTCTGTTGTTAAAGATTAATGTTCAAGCAAGAT 993
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 898 GAGCCAAATGAAGAAAGGTGATCAAGTTCAATGTGTGTGAAAGAACAGATCTCTGAGAG 957
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 994 CGGGTGACAAAGGCCATAGAGAAATCAAAATCTGCCAAGAAATTAATCTGCGAAGAGAA 1053
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 958 CGAATCCGCAAGTGGGTCAAGAGGCTGAGTAAAGAGCCGCCAAGGACACCAAGCGGCCGC 1017
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1054 CTGAGTCTCTTTTCAAGCAACTGCCACACATCAGCAGC 1094
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1018 CTGAGATGATTTCTTCAAGGTGACCGGCTCATCTCTTTCAGC 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-08-455-968E-28
; Sequence 28, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

```

NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 18985-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1144 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (polynucleotide)
US-08-455-968E-28

Query Match
Best local Similarity 56.6%; Score 302.6; DB 2; Length 1144;
Pred. No. 2.2e-80;
Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

QY 37 ATGGGCAATCAAGGTTTGAGGAACCTGCGGACATGCGCCCAAGGCGATGAGAGAG 96
DB 1 ATGGGAATTAAGGCCCTGGCCCAACTAATTTGATGTGCCCCCGACGTCATCGGGAG 60
QY 97 CAGAAGTTCAGAGCTACTTCGGCCGCAAAATGCGGTGACGCGCAGCATGACATATAC 156
DB 61 AATGACATCAAGAGCTACTTTGGCCGTAAGTGCGCATTTGATGCTCTATGACATTTAT 120
QY 157 CAGTTCCTGATTTGATGGAAGACAGCATGGAACCTCTCAAAATGAGCTGTGAA 216
DB 121 CAGTTCCTGATTTGCTGCTCCAG--GTTGGGGATGTGCTGCAGATGAGGGGTGAG 177
QY 217 GTGCTGATTTGCAAGGAATGTTCAACGCGCAATAGATTACTTGAAGCGGGAATC 276
DB 178 ACCACGACCACTGATGAGCATGTTTACCGCACTTGGCATGATGAGAAACGGCATC 237
QY 277 AAGCCAGTTTATGTTTATGCGCAAGCTCTGATATGAAGAACAAGACTTGTCTAA 336
DB 238 AAGCCGCTGATGCTTGTATGAGCAAGCCGACAGCTCAAGTCAGCGCAGCGCAAA 297
QY 337 AGATACCTCAAAAGAGTATGACACCAAGATCTGACTGACGACGATGAGGTAGAGAT 396
DB 298 CGCAGTGAAGGCGGCTGAGGCGAGAGACAGCTGACAGCTCAGCTGCTGGGGCC 357
QY 397 AAGATGCGATTTGAAATTTGAGCAAGAGACTGTAAAGTCTCAAGGCAACAACGAA 456
DB 358 GAGCAGAGAGTGAAGAAATTCATACAGCGGTGTGAAGTCTACTAAGCAGCAATGAT 417
QY 457 GATTGTAAGGCTATTAAAGCTTATGAGGCTTCTGTTGAGAGCACTTCTGAAGCA 516
DB 418 GAGTCAAAACATCTGCTGAGGCTCATGAGCATCCCTTATCTTGTATGACCCAGTGAGCA 477
QY 517 GAACAGAGATGTCAGGCTTTTGCATTAAGATTAAGTGTGCTGTTTGAAGAAT 576
DB 478 GAGGCGACGCTGCTGCTGCTGTAAGGCTGCAAAATCTATGCTGGGCTACCGAGAC 537
QY 577 ATGACACTCCCTTACTTTTGGGGCTCACGGTTCCTTCTGCTAATTAATGATCAAGTTCC 636
DB 538 ATGACACTCCCTTACTTTGCGACGCTGCTGCTATGCGAATGATGCGCAGTGAAG 597
QY 637 AAGAAATACCTGATGAGAAATTTGATGTCGCAAGSTTTTGGAGGAGCTTGAACCTACC 656
DB 598 AAAAAAGCTCAATTCAGGAATTCACCTGAGCGGATTTGCGAGGAGCTGGGCTGAAC 657
QY 697 ATGACACGATTCATTTGATGATCCTGTGATGCTGATTTGATAGATCAAA 756
DB 658 CAGGAAACGTTTGTGATCTGTGATCTCTGTAGGCACTGACTGTGAGAGATCCGG 717
QY 757 GGTATCGGGGGGCAAAACAGCTCTGAACCTTATTCGCAACATGGGCTCATGAAGATC 816
DB 718 GGTATTTGGGCGCAAGCGGCTGTGACCTCATTCAGAAAGCAGAGACATGAGAGATG 777
QY 817 TTGAGATCTTAAATGAAGATATCAATTCCTGAGGAGCTGGCCTTACCAAGAGCT 876

```

```

DB 778 GTGGGCGACTTGAACCCCAACAGTACCTGTGCGCAGAAAATTTGGCTCCACAGAGAGCT 837
QY 877 CGAGCCTTTGTTCAAGAGCCCTAATG---TCACATTTGATATTTCCGTAGCTAAATGACT 933
DB 838 CACGAGCTTTCTTGGAAACCTGAGGCTCTGTGACCCCAAGTCTGTGAGCTGAAGTGAAG 897
QY 934 GCACCTGATGAGAGGCTCTCATTAAGTTTCTGTGTAAGTAATTAATGTTTCAACGAAGAT 993
DB 898 GAGCAATTAAGAAAGAGCTATCAATGCTCATGTTGTTGTAAGAAAGCAGTTCTGTGAGAG 957
QY 994 CGGGTGAACAAGGCCATPAGAGACATCAATCTCCCAAGATTAATGCTGCGCAAGAGA 1053
DB 958 CGAATCCCGAGTGGGCTCAAGAGGCTGAGTAAGAGCCGCCAAGGACGACCCAGGCGCC 1017
QY 1054 CTCGAGTCCCTTTTCAAGCCAACTGCCACCATCAGCACC 1094
DB 1018 CTGGATGATTTCTTCAAGGTGACCGGCTCACTCTTTCAGC 1058

RESULT 7
US-08-455-968E-9
Sequence 9, Application US/08455968E
Patent No. 5874283
GENERAL INFORMATION:
APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chin-Lin
APPLICANT: Lieber, Michael
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2033 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 104..1237
US-08-455-968E-9

Query Match
Best local Similarity 20.0%; Score 276.6; DB 2; Length 2033;
Pred. No. 1.7e-72;
Matches 638; Conservative 0; Mismatches 479; Indels 24; Gaps 5;

QY 37 ATGGGCAATCAAGGTTTGAGGAACCTGCTGGCGGACATGCGCCCAAGGCGATGAGAGAG 96
DB 104 ATGGAATTCACGGCTTCCCAAACTAATTTGCTGATGTGGCCCCCAAGTGCATCCGTAG 163
QY 97 CAGAATTTGAGAGCTACTTGGCGCCCAAAATTCGCGCTGACGCGCAGCATGACATATAC 156

```

```

Db      164  AATGACATCAAGAGTACTTGGTGCACAAAGGCCATGAGCTCCATGATGAGCATCTAC 223
QY      157  CAGTTCTCTGATTGTATGTTGAAGAGCAGCATGAACTCTCACAATGAAAGCTGTGAA 216
Db      224  CAGTTCTCTGATTGTCTTGTCTAG---GGTGGGATGTGCTGCACAAAGCAGAGAGGTGAG 280
QY      217  GTCATAGTATGATTGCAAGAGTGTTCACACCGACATAGATATACATGTAAGCCGGATC 276
Db      281  ACCACAGGCC---TGATGGGCAATGTTCTACCTACCATGCG---CATGGACAATGGCATC 334
QY      277  AAGCCAGTTATGTTTGTATGAGCAAGCTCCTGATATGAGAAACAAGAGCTTGCTAAA 336
Db      335  AAGCTGTGTGATCTCTTGTATGCAAAACACACAGCTGAACTCAGGCGGAGCTGCCAAG 394
QY      337  AGATAGTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGGAGAGTATGAGATG 396
Db      395  CGCAGTGAAGAGCGCGCGCGAGGCTGAGAAAGCACTGACAGAGGCTCAGCAGGCTGGATG 454
QY      397  AAGATGTCGATTAAAAAATTGAGCAAGAGCATGTAAGGCTCACAAGSCAACACACGAA 456
Db      455  GAGGAGAGAGGTGAGAAAGTTCACCAAGAGGCTGGAAGGCTCACCAACACACATGAT 514
QY      457  GATTGTAAACGGCTATTAAAGACTTATGAGGAGTTCCTGTTGTAGAGGAGCTTCTGAACA 516
Db      515  GAGTGAACAACTGCTGAGCCCTCATGGGCATCCCTTACCTTGATGACACCGAGGAGCA 574
QY      517  GAAGCAAGATGTGACAGCTTGTGCAATAAACGATTAAGGTGTGCTGCTTGCTGAGAGAT 576
Db      575  GAGGCGACAGTGTGCTGCTGCAAAAGGCTGCAAAAGTCTATGTCGCGGCAACGAGAGAC 634
QY      577  ATGAGTCCCTTACTTTTGGGAGCTCACAGGTTCTTCTGCTTATGATGATCAAGTTC 636
Db      635  ATGAGTCCCTTACTTTTGGGAGCTCACAGGTTCTTCTGCTTATGATGATCAAGTTC 694
QY      637  AAGCAATATCTGTGATGCAATTTGATGTGGCAAGTTTGGAGAGCTTGAAGTCAAC 696
Db      695  AAGAGGTGCGCCATCCAAAGTTCATGAGACCGCGTCCGACAGGAGCTGGGTCTGTAAC 754
QY      697  ATGAGACAGTTCATTTATGATGTCATCTGATGATGATGATGATGATGATGATGATGAT 756
Db      755  CAGGAGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 814
QY      757  GGTATGGGGGGGCAACAGCTGGAACCTTATTCGATCAACATGAGTGGTCCATGAAAGCATC 816
Db      815  GGCATTTGGCCCAAGCGGCTGTGATTCATCAGAAACATMAAGCATGACGAGAGATC 874
QY      817  TTGAGATCTTTAATAAGCAGATATCAATTCCTGAGGAGTGGCTTACCAGAAGCT 876
Db      875  GTGAGGAGGTGAGACCCCAAGCAAGTACCCGTTCCAGAGAACTGGCTCCACAGAGAAC 934
QY      877  CGAGCTTGTTCAGAGCCTTA---TGTCACTTGGATATTCCTGAGCTAAATGAGCT 933
Db      935  CACACAGCTCTTCTGAGCAGAGAGTGTGAGACCAAGTGTGTGAGACTGAAGGAGGAGC 994
QY      934  GCACCTGATGAGAGGAGGTCTCATTAAGTTCCTGATAAAGATTAATGTTTCAACAGAT 993
Db      995  GAGCCAAATGTAAGAGAGTGTGCAATTTATGTGTGTAAGAAACAGACTTTTTCAGAG 1054
QY      994  CGGGTGCACAAAGCCATAGAGAAATCAATTCGCCAAGATTAATTCCTCGCAGAGAA 1053
Db      1055  CGAATTCGCGGTGGGTCAAGCGGCTGATTAAGACCGGACAGGAGCAGCACCAGGAGGC 1114
QY      1054  CTGAGATCTCTTTTCAAGCAGCTGCCACACATGAGCAGCTTAAACGGAAGAGACT 1113
Db      1115  CTGAGATGATTCTTCAAGGTGACAGGCTCAGCTCTCTCAGC-----TAAAGGC 1152
QY      1114  TCGGATTAACAAGCAAGCAGCTGCGAACAAGAAACAAGGCTGTGTGAAGAAGAAA 1173
Db      1163  AAGGAGCCAAAGCCCAAGGGGCTCTAGAGAAAGCAAGACTGTGGGAGCGGCGAAG 1222
QY      1174  T 1174

```

```

Db      1223  T 1223

RESULT 8
US-08-455-968E-4
; Sequence 4, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1930 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-455-968E-4

Query Match      19.7%; Score 272.2; DB 2; Length 1930;
Best Local Similarity 55.1%; Pred.No. 3,4e-71;
Matches 629; Conservative 0; Mismatches 488; Indels 24; Gaps 4;

QY      37  ATGGGATCAAGGCTTTGACGAAGTCTGGCGGACAAATCGCGCGACAGCCAGCATGAGAG 96
Db      1  ATGGGAATTCAGGGCTTGCCAAACTAATTCGTATGTGGCCCCAGTGCATCCGTGAG 60
QY      97  CAGAAATGTAGAGTACTTGGCGGCAAAATCGCGCGACAGCCAGCATGAGCATATAC 156
Db      61  AATGACATCAAGAGCTACTTGTGTGTAAGTGGCATCAATGCCCTCATGAGCATCTAC 120
QY      157  CAGTTCTCTATTTGTAGTGAAGGACAGCATGAAACTCTCACAAATGAAGTGTGAA 216
Db      121  CAGTTCTCTATTTGTGTGTCAG---GGTGGGATGTGCTGCAAGAACGAGAGGAGT 177
QY      217  GTCATAGTATGTTGCAAGGATGTTCAACCGACAAATAGATTAACGGAAGCGGATC 276
Db      178  ACCACAGGCTCATGAGCATGTTAT-----GGCAACCATCCGATGAGAAATGGCATC 231
QY      277  AAGCCAGTTATGTTTGTATGAGCAAGCTCTCTGTATGAAGAAACAAGAGCTTGCTAAA 336
Db      232  AAGCTGTGTAGCTTTGATGCAAAACACACAGCTGAGAGTGAAGCGGTGGCCAG 291
QY      337  AGATAGTCAAAAAGAGATGATCAACCAAGATCTGACTGAGGAGTATGAGAGTATGAGAT 396
Db      292  CGCAGTGAAGAGCGCGCGAGGCTGAGAAAGCACTGACAGAGGCTCAGAGAGGCTGGAGT 351

```

```

OY 397 AAAGATCGATGAAAAATGACAGAGAGACTGTAAGAGTCAACAAGCAACACGAA 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 GAGAGAGAGGTGAGAGAGATTGACCAAGAGAGCTGTGAAGGTACCAAGCAACATATAT 411
OY 457 GATTGTAACGGCTATTAGACTTATGGGGTCTCTGTTAGAGGACCTTTGAGCA 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 412 GAGTGCAGAACCTCGTAGCCCTCATGAGCATCCCTTACCTTGATGACCCAGCAGCA 471
OY 517 GAACGAGATGTCAGCCCTTTCATTAAGATAGGTGTGCTGCTGCTCAGAGAT 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 472 GAGGCCAGCTGTGCTGCCGGAAGAGCTGGCAAGTGTATGCTGCGGCCAGGAGG 531
OY 577 ATGAGACTCCTTACTTTTGGGCTCCAGGTTCTTCGATTTAATGATCCAGTTCC 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 ATGAGACTCCTTACTTTTGGGCTCCAGGTTCTTCGATTTAATGATCCAGTTCC 591
OY 637 AAGAAATATCTGTGATGATTTGATTTGCAAGTTTGGAGAGTTGAACTCAC 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 AAGAGAGTGCCTATCCAGAGTTCCATGAGAGCGGCTCTCGAGAGCTGGGTGTGAC 651
OY 697 ATGACACAGTTTCATGATTTTGTGATCTGTGTGATGATGATGATGATGATGAT 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 CAGGAGAGTTTGTGATGATCTGTGATCTGTGATGATGATGATGATGATGATGAT 711
OY 757 GGTATCGGGGGGCAACAGCTCTGAAACTTATTCGACATGAGGTCATGAAAGCATC 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 712 GGCATTGGCCCAAGCGGCTGTGATCTCATTCAGAGATTAAGAGATGAGAGAT 771
OY 817 TTGAGATCTTAAATAAGACAGATATCAATTCCTGAGAGAGTGCCTTACCAAGACCT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 772 GTGAGCGGCTGTGAGAGCCCAAGATGATCCCGGTTCAGAGAGATGCTCCACAGAGAGC 831
OY 877 CGAGCGTGTTCAGAGAGCTTAA--TGTCAATTTGATTTCTGAGAGTAAATGAGT 933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 832 CAGAGAGTCTTCTCGAGAGCCAGAGAGTGTGAGAGCCAGAGTCTGTGAGAGTGTGAG 891
OY 934 GCACCTGATGAGAGAGGCTCATATGTTCTGTGTAAGAGATATGATGATGATGATGAT 993
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 892 GAGCAATTAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
OY 994 CGGATGCAAGAGCCATAGAGATCAATTCGCCAAGATTAATTCGCCAAGAGAGA 1053
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 952 CGATTGCGAGTGGGCTGAGAGCGGCTGATGATGAGAGCGCCAGAGAGAGAGAGAG 1011
OY 1054 CTGAGTCTTTTCAAGCCACTGCCACACATCAGACCCGTAAACGGAAGAGACT 1113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1012 CTCGATGATTTTCAAGGTCAGAGGCTCACTCTCTCAGC-----TAAAGGC 1059
OY 1114 TCGGATAAACAGAGAGAGCTGCGAAGCAAGAAACAGAGCTGTGAAAGAGAA 1173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1060 AAGGAGCCAAACCCAGAGGCTGCTAAGAGAAAGCAAGAGCTGGGAGAGGAGAG 1119
OY 1174 T 1174
Db 1120 T 1120

```

## RESULT 9

US-08-455-968E-6  
 ; Sequence 6, Application US/08455968E  
 ; Patent No. 5874283

## ; GENERAL INFORMATION:

; APPLICANT: Harrington, John L.  
 ; APPLICANT: Hsieh, Chih-Lin  
 ; APPLICANT: Lieber, Michael  
 ; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease  
 ; NUMBER OF SEQUENCES: 63  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, 8th floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA

```

; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-455-968E-6

```

```

Query Match      19.6%; Score 270.2; DB Z; Length 1149;
Best Local Similarity 55.5%; Pred. No. 1e-70;
Matches 599; Conservative 0; Mismatches 453; Indels 27; Gaps 3;

```

```

OY 37 ATGGGATGATGAGGTTTGGAGAACTGCTGGCGGACATGCGCCCAAGCGATGAGAG 96
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGGGATTAAGAGTTTGAATGCATTTATATGGAACATGTTCCCTGCTATGAGAA 60
OY 97 CACAAGTTGAGAGCTACTTCCGCCGCAAAATCCGCTGACGCGCAGATGATGATATAC 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AGCGATTAAGAGCTTTTGGCAGAAAGGTTGGCATGATGCTCTATCTCTATAT 120
OY 157 CAGTTCTGATTTAGTTTGAAGAGAGAGCATGAAACTCTCAAAATGAAAGCTGTGAA 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CAGTTTAAATGCTGTAGACAGCAAGAGCGGTGGAGTTGACCAATGAAAGCGGTGAA 180
OY 217 GTACATGATCATTTGCAAGAGATGTTCAACCGGACAAATTAAGATTAAGAGCGGATC 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 ACAAGCTCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
OY 277 AAGCGATTTATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 AAGCTTTTATGCTTTCGACGCAAACTCCAGCTTGAAATTCATGATGATGACAAAG 300
OY 337 AGATCTCAAAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 CGGTCTTCAAGAGGCTGGAACAGAAACAGAAACAGAGAGCAACAGAAATGGA 360
OY 397 AAGATGCGATTGAAATAATTGAGCAAGAGACTGTAAGGTCACAGAGCAACAGAGAA 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 AAGATGAGCAAGAAAG-----AGATTGTGAAGGCTGTCAAAAGACATATATGAA 411
OY 457 GATTGTAACGGCTATTAGACTTATGGGGTCTCTGTTTGAAGAGCACTTTGAAAGCA 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 412 GAAGCCCAAAATTAAGACTAATGGGAATCCATATATATATAGCCCAACAGAGCT 471
OY 517 GAAGCAAGATGTCAGCCCTTTGCAATAAGCATATAGAGGTGCTGCTGCTTCAAGAT 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 472 GAGGTCATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531
OY 577 ATGAGTCCCTTACTTTTGGGCTCCAGGTTCTCTGCTCATTTAATGATGATGATGAT 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 ATGAGCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
OY 637 AAGAAATACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 AAGAGAGAACCGATTGACGAATATGATGATGATGATGATGATGATGATGATGATGAT 651

```







```

; APPLICANT: Kosugi, Yoshitsugu
; APPLICANT: Matsui, Eriko
; APPLICANT: Kawasaki, Satoko
; TITLE OF INVENTION: THERMOSTABLE FLAP ENDONUCLEASE DERIVED FROM A
; TITLE OF INVENTION: HYPERTHERMOPHILE BACTERIUM BELONGING TO THE GENUS
; TITLE OF INVENTION: PYROCOCUS
; FILE REFERENCE: 081356-0121
; CURRENT APPLICATION NUMBER: US/09/146,319
; EARLIER APPLICATION NUMBER: JP 239440/1997
; EARLIER FILING DATE: 1997-09-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1032
; TYPE: DNA
; ORGANISM: Pyrococcus horikoshii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
; US-09-146-319-1

```

```

Query Match          9.8%; Score 135.6; DB 4; Length 1032;
Best Local Similarity 52.1%; Pred. No. 1.2e-30;
Matches 303; Conservative 0; Mismatches 279; Indels 0; Gaps 0;

QY 86 CGATGAGGAGCAGAGAGTTGAGAGCTACTCGGCCGCAAAATCGCGTCGACGCCAGCA 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29 CGAGGAGGAGATAGATCTTGAATACTGTATGGAAGAGATGACATATGATCCCTAA 88
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 146 TGAGCATATACAGTTCCTGATTTAGTTGGAAGAGACAGGATGGAATCTGCACAATG 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 89 AGCCCATCTATCAGTTTATCAAGATAGACAGAGGATGGAACCACTTATGAGACT 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 206 AGCTGTGAGAGTCACTGATCTTTGCAAGATGTTCAACCGGACATTAAGTTACTGG 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 149 CTAAAGGTAGGATACCTCTCTATTAGTGGCTCTTTTATAGAACGTAATTAATGG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 AAGCGGAGATCAAGCAGTTTATGTTTGTGATGCAAGCCTCTGATATGGAAGAAAG 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 209 AAGCGGATTTAAGCGGCGCTACGCTTTGATGGAAGACCTCCGGAATTCAGAAAGAG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326 AGCTTCTAAAGATATCTCAAAAAGAGATGATGCAACCAAGATCTGACGTAGGCACTAG 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 269 AGCTCGAAAAAGAGAGAGAGTAGAGAAAGAGCAAACTAAATGGAAGAAAGCTCTAG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386 AGGTAGAGATTAAGATGCGATTGAAAAATTGAGCAAGAGACTGTAAAGGTCAAGAGC 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 329 CCAAGGAAACCTGGAGAGAGCTAGAGAAATACGTCGAAGGGCAACTAAGTTAATGAA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 446 AACACAGCAGAGATTTGTAACGGCTATTAGACTTATGGGGGTTCTGTTAGAGGCAC 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 TGCTAATCGAAGATGCAAGAAAGCTTTTGCAACTAATGGGAATACCATTAATTCAGCCTC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 506 CTTCAGAGCAGAGCAGAGATGTCAGCCCTTTGCATTAACGATTAAGTGTTCCTGTTG 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 CAAGTGAAGAGAGAGAGAGAGAGGATACATGCGCAAGTAAAGGGATGTCTACGCGTCAG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 566 CTTCAGAGATATGATGATCCCTTACTTTTGGGCTCCACGGTTCCTTCGTCATTTAATGG 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 509 CGAGTCAAGATTAATGATTAATCTACTACTCTTTGGTCTCCAAAGGTGATTAAGAAATCTGACAA 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 626 ATCCAAGTTCCAGAAATAATACCTGTGATGAATTTGATGTTG 667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 TTACGGGAAAAAGAAAGATGCTGGGAAAAAGATGTTTACGTTG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15  
US-09-175-973-1

; Sequence 1, Application US/09175973A  
; Patent No. 6255081  
; GENERAL INFORMATION:  
; APPLICANT: MATSUI, Ikuno

```

; APPLICANT: ISHIKAWA, Kazuhiko
; APPLICANT: KOSUGI, Yoshitsugu
; APPLICANT: MATSUI, Eriko
; APPLICANT: KAWASAKI, Satoko
; TITLE OF INVENTION: THERMOSTABLE FLAP ENDONUCLEASE DERIVED FROM A
; TITLE OF INVENTION: HYPERTHERMOPHILE BACTERIUM BELONGING TO THE GENUS
; TITLE OF INVENTION: PYROCOCUS
; FILE REFERENCE: 081356/0126
; CURRENT APPLICATION NUMBER: US/09/175,973A
; EARLIER APPLICATION NUMBER: JP 239440/1997
; EARLIER FILING DATE: 1997-09-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1032
; TYPE: DNA
; ORGANISM: Pyrococcus horikoshii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
; US-09-175-973-1

```

```

Query Match          9.8%; Score 135.6; DB 4; Length 1032;
Best Local Similarity 52.1%; Pred. No. 1.2e-30;
Matches 303; Conservative 0; Mismatches 279; Indels 0; Gaps 0;

QY 86 CGATGAGGAGCAGAGAGTTGAGAGCTACTCGGCCGCAAAATCGCGTCGACGCCAGCA 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29 CGAGGAGGAGATAGATCTTGAATACTGTATGGAAGAGATGACATATGATCCCTAA 88
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 146 TGAGCATATACAGTTCCTGATTTAGTTGGAAGAGACAGGATGGAATCTGCACAATG 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 89 AGCCCATCTATCAGTTTATCAAGATAGACAGAGGATGGAACCACTTATGAGACT 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 206 AGCTGTGAGAGTCACTGATCTTTGCAAGATGTTCAACCGGACATTAAGTTACTGG 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 149 CTAAAGGTAGGATACCTCTCTATTAGTGGCTCTTTTATAGAACGTAATTAATGG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 AAGCGGAGATCAAGCAGTTTATGTTTGTGATGCAAGCCTCTGATATGGAAGAAAG 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 209 AAGCGGATTTAAGCGGCGCTACGCTTTGATGGAAGACCTCCGGAATTCAGAAAGAG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326 AGCTTCTAAAGATATCTCAAAAAGAGATGATGCAACCAAGATCTGACGTAGGCACTAG 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 269 AGCTCGAAAAAGAGAGAGAGTAGAGAAAGAGCAAACTAAATGGAAGAAAGCTCTAG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386 AGGTAGAGATTAAGATGCGATTGAAAAATTGAGCAAGAGACTGTAAAGGTCAAGAGC 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 329 CCAAGGAAACCTGGAGAGAGCTAGAGAAATACGTCGAAGGGCAACTAAGTTAATGAA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 446 AACACAGCAGAGATTTGTAACGGCTATTAGACTTATGGGGGTTCTGTTAGAGGCAC 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 TGCTAATCGAAGATGCAAGAAAGCTTTTGCAACTAATGGGAATACCATTAATTCAGCCTC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 506 CTTCAGAGCAGAGCAGAGATGTCAGCCCTTTGCATTAACGATTAAGTGTTCCTGTTG 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 CAAGTGAAGAGAGAGAGAGAGGATACATGCGCAAGTAAAGGGATGTCTACGCGTCAG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 566 CTTCAGAGATATGATGATCCCTTACTTTTGGGCTCCACGGTTCCTTCGTCATTTAATGG 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 509 CGAGTCAAGATTAATGATTAATCTACTACTCTTTGGTCTCCAAAGGTGATTAAGAAATCTGACAA 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 626 ATCCAAGTTCCAGAAATAATACCTGTGATGAATTTGATGTTG 667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 TTACGGGAAAAAGAAAGATGCTGGGAAAAAGATGTTTACGTTG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: November 5, 2002, 11:03:04  
Job time : 44.2033 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 10:28:52 ; Search time 1181.61 seconds  
15774.429 Million cell updates/sec

Title: US-09-805-311-5  
Perfect score: 1381  
Sequence: 1 cgaccacgcgcgcgcgcac.....ltgaaaaaaaaaaaaaa 1381

Scoring table: IDENTITY\_NTC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
EST:  
1: em\_estda:  
2: em\_esthum:  
3: em\_estin:  
4: em\_estnu:  
5: em\_estov:  
6: em\_estpl:  
7: em\_estro:  
8: em\_hlc:  
9: gb\_estl:  
10: gb\_estc2:  
11: gb\_hlc:  
12: gb\_gss:  
13: em\_gss\_hum:  
14: em\_gss\_iny:  
15: em\_gss\_pln:  
16: em\_gss\_vit:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	770.2	55.8	901	10	BG837708
2	523.8	37.9	550	10	BE639422
3	511.2	37.0	554	10	BE639421
4	509	36.9	586	9	AI881599
5	501	36.3	532	9	AM000375
6	497.8	36.0	553	9	AM562789
7	465.2	33.7	470	9	AI861468
8	463.4	33.6	474	9	AM539173
9	458.2	33.2	468	9	AV913663
10	448	32.4	467	9	BE186786
11	420.4	30.4	733	9	AM774700
12	411.6	29.8	539	10	BC263333
13	409.8	29.7	475	9	AM562517
14	402.2	29.1	414	9	AM288784
15	397	28.7	456	9	AI065689
16	393	28.5	553	9	AI834484
17	378.8	27.4	632	9	AM585913

18	372.6	27.0	609	10	BG096545
19	370.4	26.8	734	10	BE416584
20	369.6	26.8	820	10	BG414505
21	364.6	26.4	645	9	AM691681
22	348	25.2	460	10	BE417817
23	347.2	25.1	683	9	AV918966
24	346.8	25.1	454	9	AV915644
25	325.4	23.6	539	10	BP098320
26	321.8	23.3	363	9	AM562518
27	317.2	23.0	470	9	AM288831
28	314.8	22.8	376	9	AI065546
29	314.8	22.8	579	9	AV833393
30	314.4	22.8	592	9	AI780966
31	308	22.3	526	10	BI419720
32	300.6	21.8	695	9	BE052579
33	273.8	19.8	942	9	AL560007
34	273	19.8	549	9	AM696665
35	265.6	19.2	903	9	AL531350
36	263.8	19.1	851	9	AL519300
37	248.6	18.0	689	10	BE311755
38	245.4	17.8	769	10	BG574950
39	244.4	17.7	660	10	BJ011893
40	244.2	17.7	806	9	AL560395
41	237.6	17.2	762	10	BG337603
42	236	17.1	866	10	BI827898
43	234.4	17.0	785	10	BG756459
44	233.6	16.9	704	9	AM783795
45	233.4	16.9	963	12	CNS06WXR

## ALIGNMENTS

RESULT 1  
LOCUS BG837708 901 bp mRNA linear EST 25-MAY-2001  
DEFINITION Zm10\_01f08.A Zm10-AAFC\_ECORC\_Fusarium\_graminearum\_corn\_silk Zea  
mays cDNA clone Zm10\_01f08, mRNA sequence.  
ACCESSION BG837708  
VERSION BG837708.1 GI:14204031  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
AUTHORS Harris,L.J., Balcerzak,M., Allard,S., Sapano,A., Couroux,P., De Moors,A., Hattori,J.I., Quellet,T., Robert,L.S., Singh,J.A., Sprott,D. and Tinker,N.A.  
TITLE Expressed Sequence Tags from Maize Silk Six Hours After Silk  
JOURNAL Channel inoculation with Fusarium graminearum  
COMMENT Unpublished (2001)  
CONTACT: Harris, Linda J.  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-Food Canada  
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
CANADA  
Tel: (613) 759-1314  
Fax: (613) 759-6566  
Email: harris.lj@em.agr.ca.

## FEATURES

source  
Location/Qualifiers  
1..901  
/organism="Zea mays"  
/cultivar="CO388"  
/db\_xref="taxon:4577"  
/clone="Zm10\_01f08"  
/clone\_lib="Zm10-AAFC\_ECORC\_Fusarium\_graminearum\_corn\_silk"  
/tissue\_type="Silk"  
/dev\_stage="4-5 days post-silk emergence"  
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI; Site 2: XhoI; Field-grown corn was silk channel-inoculated

In the morning (~10 am) with 1 ml of a macroconidial suspension (500,000 spores/ml) of *Fusarium graminearum* and silk channels were collected and immediately frozen in liquid nitrogen 6 hours later. RNA was extracted from silk tissue between 1 cm below and above the inoculation point in the silk channel, RNA from five silk channels was pooled."

BASE COUNT 267 a 174 c 208 g 249 t 3 others  
ORIGIN

Query Match 55.8%; Score 770.2; DB 10; Length 901;  
Best Local Similarity 98.3%; Pred. No. 1.2e-153;  
Matches 797; Conservative 2; Mismatches 10; Indels 2; Gaps 2;

```

QY 557 TCCTGTTGCTTCAGAAATATGAGTCCCTTACTTTGGGGCTCCAGGTTCTCTGTC 616
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 TCCTGTTGCTTCAGAAATATGAGTCCCTTACTTTGGGGCTCCAGGTTCTCTGTC 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 617 ATTTAATGATCCAGTCCAGAAATATCTGTATGAAATTTGATTTGSCCAAGGTTT 676
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 ATTTAATGATCCAGTCCAGAAATATCTGTATGAAATTTGATTTGSCCAAGGTTT 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 677 TGGAGAGCTTGAACCTCACCATGACACCTTATGATTGTCATCTCTGTGTGATGTG 736
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 TGGAGAGCTTGAACCTCACCATGACACCTTATGATTGTCATCTCTGTGTGATGTG 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 737 ACTATGTGATAGCATCAAGATATCGGGGGCAACAGCTCTGAAACTTATTCGTACAC 796
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 ACTATGTGATAGCATCAAGATATCGGGGGCAACAGCTCTGAAACTTATTCGTACAC 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 797 ATGGGTCCATAGAAAGCATTTGGAGATCTTAATTAAGACAGATATCAATTTCC-TCAG 855
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 ATGGGTCCATAGAAAGCATTTGGAGATCTTAATTAAGACAGATATCAATTTCC-TCAG 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 856 GACTGGCTTACC-AAGAACTCTGACGCTTGTTCAGAGAGCCCTAATGTACATTTGATAT 914
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 GACTGGCTTACC-AAGAACTCTGACGCTTGTTCAGAGAGCCCTAATGTACATTTGATAT 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 915 TCCGAGCTTAATGAGCTGACCTGATGAGAGGGTCCATTAATTTCCCTGTTAAAGA 974
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 TCCGAGCTTAATGAGCTGACCTGATGAGAGGGTCCATTAATTTCCCTGTTAAAGA 420
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 975 TAATGTTCACGAAAGATGGGTGACAAAGCCATAGAGAGATCAAAATTCGCCAAGAA 1034
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 421 TAATGTTCACGAAAGATGGGTGACAAAGCCATAGAGAGATCAAAATTCGCCAAGAA 480
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1035 TAAATGTCGCAAGGAAGACTCGAGTCTTTTCAAGCCAACTGCGCACCATCAGACG 1094
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 481 TAAATGTCGCAAGGAAGACTCGAGTCTTTTCAAGCCAACTGCGCACCATCAGACG 540
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1095 GCTAAACGAGAGAGACTTGGATTAACCAAGCAAGGAGCTGCGAACAAGAAACAAA 1154
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 541 GCTAAACGAGAGAGACTTGGATTAACCAAGCAAGGAGCTGCGAACAAGAAACAAA 600
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1155 GCGTGTGGAAGAGAAATATCTTGATGCTGTGATGACAACTACAGACTACGAAAGCA 1214
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 601 GCGTGTGGAAGAGAAATATCTTGATGCTGTGATGACAACTACAGACTACGAAAGCA 660
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1215 GCGGTGCGTGATCACTTGGCTTAGATTAATTAACCTCCCTGTTTAACTCAGAGCTTTGG 1274
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 661 GCGGTGCGTGATCACTTGGCTTAGATTAATTAACCTCCCTGTTTAACTCAGAGCTTTGG 720
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1275 TAAAGTTGCGCATGTTTCAAGCTGGGGTAAGTTAGTTTGGAGAGAAATTTGGGTA 1334
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 721 TGAAGTTTGGCCATGTTTCAAGCTGGGGTAAGTTAGTTTGGAGAGAAATTTGGGTA 780
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1335 CCAAGTAACAAACTTATCGCTGTTTTTTGA 1365
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 781 CCAAGTAACAAACTTATCGCTGTTTTTTGA 811
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

RESULT 2  
BE639422

LOCUS BE639422 550 bp mRNA linear EST 30-AUG-2000  
DEFINITION 946033A02.y2 946 - tassels primordium prepared by Schmidt Lab Zea  
ACCESSION mays CDNA, mRNA sequence.  
VERSION BE639422  
KEYWORDS BE639422.1 GI:9952839  
SOURCE EST.  
ORGANISM Zea mays.  
Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 550)

AUTHORS Walbot,V.

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL Unpublished (1999)

COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 946033 row: A column: 02.

FEATURES  
source  
1..550  
/organism="Zea mays"  
/cultivar="OH43"  
/db\_xref="taxon:457"  
/clone\_lib="946 - tassels primordium prepared by Schmidt  
lab"  
/tissue\_type="tassels"  
/dev\_stage="just after the transition from vegetative to  
inflorescence development"  
/lab\_host="X10LR"  
/note="Organ: tassels; Vector: HybridAP; site\_1: EcoRI;  
site\_2: XhoI; George Chuck dissected immature tassels  
between imm and 3mm. Sharon Stanfield prepared the cDNA  
library in HybridAP. Sample insert size range was 350 bp  
to 3 Kb with a 1 Kb average."

BASE COUNT 177 a 116 c 149 g 108 t  
ORIGIN

Query Match 37.9%; Score 523.8; DB 10; Length 550;  
Best Local Similarity 98.5%; Pred. No. 2.9e-101;  
Matches 539; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```

QY 1 CGACCCAGCGCGTCGCGCACAGCGCGCCGACAGCAGATGGCGCATCAAGGTTTGACGAAA 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 4 CGACCCAGCGCGTCGCGCACAGCGCGCCGACAGCAGATGGCGCATCAAGGTTTGACGAAA 63
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 CTGCTGGCGGCAATGCGCCCAAGGCGATGAAGGAGCAGAGATTGAGAGCTTCTGCGC 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 64 CTGCTGGCGGCAATGCGCCCAAGGCGATGAAGGAGCAGAGATTGAGAGCTTCTGCGC 123
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 CGCAAAATCGCGCTGCGAGCGCCAGCATGAGATATACCATTTCTGATTTGTTGGAAG 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 124 CGCAAAATCGCGCTGCGAGCGCCAGCATGAGATATACCATTTCTGATTTGTTGGAAG 183
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 181 ACAGG-CATGGAATCTCTCAAAATGAAGCTGTGAAGTCACTAGTCAATTTGCAAGAAAT 239
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 184 ACTGTCATGGAATCTCTCAAAATGAAGCTGTGAAGTCACTAGTCAATTTGCAAGAAAT 243
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 240 GTTCAACCGGACATATAGATTACTGAGAGCGGGAATCAACAGCGTTTATGTTTGGATGG 299
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 244 GTTCAACCGGACATATAGATTACTGAGAGCGGGAATCAACAGCGTTTATGTTTGGATGG 303
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 300 CAAGCTCTCTGATATGAAGAAACAAGAGCTTGTCTAAAGATCTCTCAAAAAGAGATGATGC 359
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 304 CAAGCTCTCTGATATGAAGAAACAAGAGCTTGTCTAAAGATCTCTCAAAAAGAGATGATGC 363
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 360 AACCAAGATCTGACGTGAGCCTAGAGGTAGAGATAAGATGGATGAAAAATTTGAG 419
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

Db 364 AACCAAGATCTGACTGAGCAGTAGAGATTAAGATGCGATTGAAATTTGAG 423  
QY 420 CAAGAGACTGTAAGGTCACAGCAACAGCAAGATTTGTAAGCGCTATTAGACT 479  
|||||  
Db 424 CAAGAGACTGTAAGGTCACAGCAACAGCAAGATTTGTAAGCGCTATTAGACT 483  
QY 480 TATGGGGTTCCTGTTGAGAGCAGCTTCTGAGCAGAGAGATGTGACGCCCTTTG 539  
|||||  
Db 484 TATGGGGTTCCTGTTGAGAGCAGCTTCTGAGCAGAGAGATGTGACGCCCTTTG 543  
QY 540 CATTAAC 546  
|||||  
Db 544 CATTAAC 550

RESULT 3  
BE639421/c 554 bp mRNA linear EST 30-AUG-2000  
LOCUS 946033A02.x2 946 - tassal primordium prepared by Schmidt lab Zea  
DEFINITION mays cDNA, mRNA sequence.  
ACCESSION BE639421  
VERSION BE639421.1 GI:9952838  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoidae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 554)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY University  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 946033 row: A column: 02.  
Location/Qualifiers  
FEATURES  
source 1..554  
/organism="Zea mays"  
/cultivar="Oh43"  
/db\_xref="taxon:4577"  
/clone\_lib="946 - tassal primordium prepared by Schmidt  
lab"  
/tissue\_type="tassels"  
/dev\_stage="just after the transition from vegetative to  
inflorescence development"  
/lab\_host="XLOLR"  
/note="Organ: tassels; Vector: HybridAP; Site\_1: EcoRI;  
Site\_2: XhoI; George Chuck dissected immature tassels  
between 1mm and 3mm. Sharon Stanfield prepared the cDNA  
library in HybridAP. Sample insert size range was 350 bp  
to 3 kb with a 1 kb average."

BASE COUNT 145 a 131 c 110 g 168 t  
ORIGIN  
Query Match 37.0%; Score 511.2; DB 10; Length 554;  
Best Local Similarity 99.4%; Pred. No. 1.4e-98;  
Matches 513; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 848 TTCTGTGAGCGCTTCAAGAGCTGAGCGCTGTGTAAGAGCACTAATGTCATAT 907  
|||||  
Db 554 TTCTGTGAGCGCTTCAAGAGCTGAGCGCTGTGTAAGAGCACTAATGTCATAT 495  
|||||  
QY 908 TGGATATCTCTGAGTAAATGAGCTGACCTGATGAGAGGCTCTCATAGTTTCTTG 967  
|||||  
Db 494 TGGATATCTCTGAGTAAATGAGCTGACCTGATGAGAGGCTCTCATAGTTTCTTG 435  
|||||  
QY 968 TAAAGATATGTTTCAACGAAGATCGGCTACAAAGCCATTAAGAAATCAATCTG 1027

Db 434 TAAAGATATGTTTCAACGAAGATCGGTTGACAAAGCCATGAGAGAATCAATCTG 375  
|||||  
QY 1028 CCAAGATTAATCTGTCGAGAGAACTCGACCTCTTTTCAACCAACTGCCACAT 1087  
|||||  
Db 374 CCAAGATTAATCTGTCGAGAGAACTCGACCTCTTTTCAACCAACTGCCACAT 315  
|||||  
QY 1088 CAGCACCCTTAAAGGAGAGAGACTTCGATTAACAAAGCAAGCAGCTGCGCAAGA 1147  
|||||  
Db 314 CAGCACCCTTAAAGGAGAGAGACTTCGATTAACAAAGCAAGCAGCTGCGCAAGA 255  
|||||  
QY 1148 AAACCAAGCTCTGGAAGAAGAATATCTTGATGCTGATGTCACACTACACTAC 1207  
|||||  
Db 254 AAACCAAGCTCTGGAAGAAGAATATCTTGATGCTGATGTCACACTACACTAC 195  
|||||  
QY 1208 GAAGAGCGGTGGCGTGATCATCTCGCTTATATTATTAACCTCCGTTTACTCAGA 1267  
|||||  
Db 194 GAAGAGCGGTGGCGTGATCATCTCGCTTATATTATTAACCTCCGTTTACTCAGA 135  
|||||  
QY 1268 GCTTTGTAAAGTTTCGCCATGTTTCAAGCTGGGCTAAGTTAGTTGTTTGAAGAT 1327  
|||||  
Db 134 GCTTTGTAAAGTTTCGCCATGTTTCAAGCTGGGCTAAGTTAGTTGTTTGAAGAT 75  
|||||  
QY 1328 TGCTGTACCAAGTAAACAAACTTATCGCTGTTTTT 1363  
|||||  
Db 74 TGCTGTACCAAGTAAACAAACTTATCGCTGTTTTT 39  
|||||

RESULT 4  
A1881599 586 bp mRNA linear EST 02-FEB-2000  
LOCUS A1881599  
DEFINITION 606068G09.y1 606 - Ear tissue cDNA library from Schmidt lab Zea  
mays cDNA, mRNA sequence.  
ACCESSION A1881599  
VERSION A1881599.1 GI:5566733  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoidae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 586)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY University  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 606068 row: G column: 09.  
Location/Qualifiers  
FEATURES  
source 1..586  
/organism="Zea mays"  
/cultivar="Oh1043"  
/db\_xref="taxon:4577"  
/clone\_lib="606 - Ear tissue cDNA library from Schmidt  
lab"  
/tissue\_type="mixed"  
/dev\_stage="ear length from 0.5 cm - 2.0 cm"  
/lab\_host="XLOLR (Stratagene)"  
/note="Organ: immature ear; Vector: pBK-CMV; Site\_1: EcoRI  
; Site\_2: XhoI; Mixed ear tissue cDNA library from Schmidt  
lab"

BASE COUNT 186 a 132 c 159 g 109 t  
ORIGIN  
Query Match 36.9%; Score 509; DB 9; Length 586;  
Best Local Similarity 98.1%; Pred. No. 4.1e-98;  
Matches 515; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

[illegible]

# FEATURES source

Location/Qualifiers

1..553

/organism="Zea mays"

/cultivar="Ohio43"

/db\_xref="taxon:4577"

/clone\_id="660 - Mixed stages of anther and pollen"

/tissue\_type="whole premeiotic anthers to pollen shed"

/dev\_stage="premeiotic anthers to pollen shed"

/lab\_host="XLOLR"

/note="Organ: anthers; Vector: lambda Zap; Site\_1: EcORI; Site\_2: XhoI; Anther and pollen cDNA library. Directionally sequenced with 5' end at the EcoRI site. Created by Amie Franklin."

BASE COUNT 171 a 111 c 129 g 142 t

ORIGIN

Query Match 36.0%; Score 497.8; DB 9; Length 553;

Best Local Similarity 98.6%; Pred. No. 9.9e-96;

Matches 502; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 855 GGAAGTGGCTTACCAAGAAAGCTGACGCTTGTTCAGAGAGCCTAATGTACATTGGATAT 914

DB 4 GGATGGCTTACCAAGAAAGCTGACGCTTGTTCAGAGAGCCTAATGTACATTGGATAT 63

QY 915 TCCGAGCTTAAATGAGTGCACCTGATGAGAGGGTCTCATAGTTTCTGGTAAAGA 974

DB 64 TCCTGAGCTTAAATGAGTGCACCTGATGAGAGGGTCTCATAGTTTCTGGTAAAGA 123

QY 975 TAATGTTTACAGAGATGGGTGACAAAGGCCATAGAGATCAATCTGCCACAA 1034

DB 124 TAATGTTTCAATGAGATGGGTGACAAAGGCCATAGAGATCAATCTGCCACAA 183

QY 1035 TAAATGCTGCAGAGAAAGCTGAGTCTTTTTCAGCAACCTGCACACATCAGCACC 1094

DB 184 TAAATGCTGCAGAGAAAGCTGAGTCTTTTTCAGCAACCTGCACACATCAGCACC 243

QY 1095 GCTAAAGCGAAGAGACTTCGGATTAACAACAAGCAGCTGCGAACAAGAAACAA 1154

DB 244 GCTAAAGCGAAGAGACTTCGGATTAACAACAAGCAGCTGCGAACAAGAAACAA 303

QY 1155 GGCTGTGGAAGAAGAAATATCTGTGATGATGTACAACTAGCACTACGAAAGCA 1214

DB 304 GGCTGTGGAAGAAGAAATATCTGTGATGATGTACAACTAGCACTACGAAAGCA 363

QY 1215 GCGGTGGCTGATCACTTGGCTTAGATTTTAACTCCCTGTTTAACTCAGAGTTTGG 1274

DB 364 GCGGTGGCTGATCACTTGGCTTAGATTTTAACTCCCTGTTTAACTCAGAGTTTGG 423

QY 1275 TAAAGTTCCGCACTGTTTCAAGCTGGGTAGATTAGTGTCTTGAAGAGATTGTTGA 1334

DB 424 TGAAGTTCCGCACTGTTTCAAGCTGGGTAGATTAGTGTCTTGAAGAGATTGTTGA 483

QY 1335 CCAAGTACAAACTTATGCTCTTTT 1363

DB 484 CCAAGTACAAACTTATGCTCTTTT 512

## RESULT 7 AI861468/c

LOCUS 614014D03.xl 614 - root cDNA library from Walbot Lab Zea mays cDNA,

DEFINITION mRNA sequence.

ACCESSION AI861468 470 bp mRNA linear EST 19-JUL-1999

VERSION AI861468.1 GI:5525575

KEYWORDS EST.

SOURCE Zea mays.

ORGANISM Zea mays

REFERENCE Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae, PACC

1 (bases 1 to 470)

WALBOT.V. Walbot.V.

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford

# JOURNAL COMMENT

University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 614014 row: D column: 03.

## FEATURES source

Location/Qualifiers

1..470

/organism="Zea mays"

/cultivar="W23"

/db\_xref="taxon:4577"

/clone\_id="614 - root cDNA library from Walbot Lab"

/tissue\_type="root"

/dev\_stage="3-4 days old"

/lab\_host="XLOLR"

/note="Organ: root; Vector: pBluescriptII SK+; Site\_1: EcORI; Site\_2: XhoI; 3-4 days old root tissue from Walbot Lab (LM)"

BASE COUNT 112 a 113 c 93 g 152 t

ORIGIN

Query Match 33.7%; Score 465.2; DB 9; Length 470;

Best Local Similarity 99.4%; Pred. No. 8.4e-89;

Matches 467; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 889 AAGGACCTTAATGATCATGATGATATTCCTGAGCTAAATGAGCTGACCTGTAGAGAG 948

DB 470 AAGGACCTTAATGATCATGATGATATTCCTGAGCTAAATGAGCTGACCTGTAGAGAG 411

QY 949 GGTCATATAGTTTCTGCTGTAAGAAGATATGTTTCAACGAGATGGGTGACAAAGGCC 1008

DB 410 GGTCATATAGTTTCTGCTGTAAGAAGATATGTTTCAACGAGATGGGTGACAAAGGCC 351

QY 1009 ATAGAGAGATCAATCTGCCAAGATTAATCTGCGAAGAGACCTGAGTCTTTTTC 1068

DB 350 ATAGAGAGATCAATCTGCCAAGATTAATCTGCGAAGAGACCTGAGTCTTTTTC 291

QY 1069 AAGCCACTGCGCAGCATCATCAGACACCGCTAAACGGAAGGACTTGGATTAACACAGC 1128

DB 290 AAGCCACTGCGCAGCATCATCAGACACCGCTAAACGGAAGGACTTGGATTAACACAGC 231

QY 1129 AAGGCACTGCGCAGCAAGAAACAAAGGCTGCTGGAAGAAAGAAATATCTTGGATGCT 1188

DB 230 AAGGCACTGCGCAGCAAGAAACAAAGGCTGCTGGAAGAAAGAAATATCTTGGATGCT 171

QY 1189 GATGTACATCTGCACTGCAAGAGAGGCTGGCGTGTACTCTGCTTGAATTAATTA 1248

DB 170 GATGTACATCTGCACTGCAAGAGAGGCTGGCGTGTACTCTGCTTGAATTAATTA 111

QY 1249 CTCCTGTTTAACTCAGAGCTTTGGTAAAGTTGCCCATTGTTCAAGCTGGGCTAAGT 1308

DB 110 CTCCTGTTTAACTCAGAGCTTTGGTAAAGTTGCCCATTGTTCAAGCTGGGCTAAGT 51

QY 1309 TGGTTGTTTGAAGAGATGGGTGACCAAGTAAACAAACTTATCGCTG 1358

DB 50 TGGTTGTTTGAAGAGATGGGTGACCAAGTAAACAAACTTATCGCTG 1

## RESULT 8 AW559173/c

LOCUS 660065H06.xl 660 - Mixed stages of anther and pollen Zea mays cDNA,

DEFINITION mRNA sequence.

ACCESSION AW559173 474 bp mRNA linear EST 07-MAR-2000

VERSION AW559173.1 GI:7204640

KEYWORDS EST.

SOURCE Zea mays.

ORGANISM Zea mays

REFERENCE Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:

REFERENCE 1 (bases 1 to 474)  
 TITLE: Maize ESTs from various cDNA libraries sequenced at Stanford University  
 AUTHORS: Walbot, V.  
 JOURNAL: Unpublished (1999)  
 COMMENT: Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 660065 row: H column: 06.  
 Location/Qualifiers  
 1. 474

FEATURES  
 source  
 1. 474  
 /organism="Zea mays"  
 /cultivar="Ohio43"  
 /db\_xref="taxon:4577"  
 /clone\_lib="660 - Mixed stages of anther and pollen"  
 /tissue\_type="whole premeiotic anthers to pollen shed"  
 /dev\_stage="premeiotic anthers to pollen shed"  
 /lab\_host="X10LR"  
 /note="Organ: anthers; Vector: Lambda Zap; Site\_1: EcoRI; Site\_2: XhoI; Anther and pollen cDNA library. Directionally sequenced with 5' end at the EcoRI site. Created by Amie Franklin."

BASE COUNT 113 a 114 c 92 g 154 t 1 others  
 ORIGIN

Query Match 33.6%; Score 463.4; DB 9; Length 474;  
 Best Local Similarity 98.5%; Pred. No. 2e-88; Mismatches 7; Indels 0; Gaps 0;  
 Matches 467; Conservative 0;

QY 872 AAGCTCAGCCTGTTCAAGAGCCCTPAATGTCATATTCCTGAGCTAAATGCA 931  
 DB 474 AAGCTCAGCCTGTTCAAGAGCCCTPAATGTCATATTCCTGAGCTAAATGCA 415  
 QY 932 CTGCACCTGATGAGAGGGTCTCATAGTTTCTGTTAAAGTAATGTTTCAAGCA 991  
 DB 414 CTGCACCTGATGAGAGGGTCTCATAGTTTCTGTTAAAGTAATGTTTCAAGCA 355  
 QY 992 ATGGGGTGAACAAGCCATAGAGATCAATCTGCCAAGTAATCTGTCGAAGAA 1051  
 DB 354 ATCGGGTGAACAAGCCATAGAGATCAATCTGCCAAGTAATCTGTCGAAGAA 295  
 QY 1052 GACTCGAGTCTTTTCAAGCCAACTGCCACATCAGACCGCTAAAGCGAAGAGA 1111  
 DB 294 GACTCGAGTCTTTTCAAGCCAACTGCCACATCAGACCGCTAAAGCGAAGAGA 235  
 QY 1112 CTTCGGATTAACAAGCAAGCGACGTGCGAACAAGAAACAAAGGCTGTGGAAGAAGA 1171  
 DB 234 CTTCGGATTAACAAGCAAGCGACGTGCGAACAAGAAACAAAGGCTGTGGAAGAAGA 175  
 QY 1172 AATAATCTGGATCTTGATGTACACTACACTACAGTACGAAGCGAGGCTGATGTCAC 1231  
 DB 174 AATAATCTGGATCTTGATGTACACTACACTACAGTACGAAGCGAGGCTGATGTCAC 115  
 QY 1232 TCGCTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTTGGTAAGTCCCATGT 1291  
 DB 114 TCGCTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTTGGTAAGTCCCATGT 55  
 QY 1292 TTCAAGCTGGGGTAACTAGTGTGTTGAAGATTGGTACCAAGTAACA 1345  
 DB 54 TTCAAGCTGGGGTAACTAGTGTGTTGAAGATTGGTACCAAGTAACA 1

RESULT 9  
 AV913663 648 bp mRNA linear EST 18-JAN-2002  
 LOCUS AV913663 K. Sato unpublished cDNA library, cv. Haruna Nijo  
 DEFINITION germination shoots Hordeum vulgare subsp. vulgare cDNA clone

ACCESSION bag922n17 5', mRNA sequence.  
 VERSION AV913663.1 GI:18209440  
 KEYWORDS EST.  
 SOURCE Hordeum vulgare subsp. vulgare.  
 ORGANISM Hordeum vulgare subsp. vulgare.  
 REFERENCE Sato, K., Saito, D. and Takeda, K.  
 AUTHORS Barley EST sequencing project in NIG and Okayama Univ  
 TITLES Unpublished (2002)  
 JOURNAL Contact: Tadasu Shin-i  
 COMMENT Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@genes.nig.ac.jp.  
 Location/Qualifiers  
 1. 648  
 /organism="Hordeum vulgare subsp. vulgare"  
 /cultivar="Haruna Nijo"  
 /db\_xref="taxon:112509"  
 /clone\_lib="bag922n17"  
 /clone\_lib="K. Sato unpublished cDNA library, cv. Haruna Nijo germination shoots"  
 /tissue\_type="shoots"  
 /dev\_stage="germination"

BASE COUNT 174 a 158 c 185 g 131 t  
 ORIGIN

Query Match 33.2%; Score 458.2; DB 9; Length 648;  
 Best Local Similarity 86.0%; Pred. No. 2.6e-87; Mismatches 83; Indels 0; Gaps 0;  
 Matches 508; Conservative 0;

QY 13 CCGGCCACAGCCGCCGACAGAGATGGGATCAAGGGTTTGACGAACTGTGGCGAC 72  
 DB 58 CCGGCCACAGCCGCCGACAGAGATGGGATCAAGGGTTTGACGAACTGTGGCGAC 117  
 QY 73 AATGGCCCAAGCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132  
 DB 118 AATGGCCCAAGCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 177  
 QY 133 GTGAGCCAGCATGAGATATACGATTCCTGATTGTAGTTGAAGAGACAGCATGAA 192  
 DB 178 GTGAGCCAGCATGAGATATACGATTCCTGATTGTAGTTGAAGAGACAGCATGAA 237  
 QY 193 ACHTCCAAATGAACGCTGTGAAGTCACTAGTCAATTCGAAGAAATGTCACCGGACA 252  
 DB 238 ACCCTTACAAACGAGCCGCTGATGTACACAGTCAATTCGAAGGATGTCACCGGACA 297  
 QY 253 ATAAGATTACTGGAACCGGGAATCAAGCGTTTATGTTTGTGATGACCAAGCCTCTGAT 312  
 DB 298 ATAAGATTACTGGAACCGGGAATTAACCAAGTATGTTTGTGATGACCAAGCCTCTGAA 357  
 QY 313 ATGAAGAAAGAGAGTGTCTTAAGATATCTCAAAAAGAGATGATGACCAAGACATGTG 372  
 DB 358 ATGAAGAAAGAGAGTGTCTTAAGATATCTCAAAAAGAGATGATGACCAAGACATGTG 417  
 QY 373 ACTGAGGCACTAGAGTAGAGATGAAGATGCGATTGAAAAATTGAGCAAGAGACTGTA 432  
 DB 418 ACTGAGGCACTAGAGTAGAGATGCGATTGAAAAATTGAGCAAGAGACTGTA 477  
 QY 433 AAGGTCAAGCAAGCAACAGAGATGTAAACCGCTATTAAAGCTTAATGGGTGCTTCT 492  
 DB 478 AAGGTCAAGCAAGCAACAGATGTAAACCGCTATTAAAGCTTAATGGGTGCTTCT 537  
 QY 493 GTTGTAGAGCACTTCTTAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552  
 DB 538 GTTGTAGAGCACTTCTTGTGAAGAGAGATCACAATGAGTGGCTTGTGCAAGAGTGAAG 597

Oy	553	GTTGTCGCCTTGGCTCAGAAATATGACCTCCACTTTTGCGGCTCCA	603
Dd	598	GTTATGCTGTTCATCATCAAGATATGAGTCACTTACTTTTGAGACTTCA	648
<hr/>			
RESULT 10			
LOCUS		BE186786	467 bp mRNA linear EST 22-JUN-2000
DEFINITION		94601208.XI 946 - tassal primordium prepared by Schmidt lab Zea	
ACCESSION		mays cDNA, mRNA sequence.	
VERSION		BE186786	
KEYWORDS		BE186786.1 GI:8665970	
SOURCE		EST.	
ORGANISM		Zea mays.	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea. 1 (bases 1 to 467)	
REFERENCE		Walbot,V.	
AUTHORS		Maize ESTs from various cDNA libraries sequenced at Stanford University	
TITLE		Unpublished (1999)	
JOURNAL		Contact: Walbot V	
COMMENT		Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 946012 row: C column: 08.	
<hr/>			
FEATURES			
Source			
1..467			
/organism="Zea mays"			
/cultivar="OH43"			
/db_xref="taxon:4577"			
/clone_lib="946 - tassal primordium prepared by Schmidt lab"			
/tissue_type="tassels"			
/dev_stage="just after the transition from vegetative to inflorescence development"			
/lab_host="XILDR"			
/note="Organ: tassels; Vector: HybriZAP; Site:1; Ecobl; site:2; XhoI; George Chuck dissected immature tassels between lmm and 3mm. Sharon Stanfield prepared the cDNA library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 kb average."			
BASE COUNT			
108 a 110 c 96 g 153 t			
ORIGIN			
Query Match 32.4%; Score 448; DB 9; Length 467;			
Best Local Similarity 98.7%; Pred. No.3.8e-85;			
Matches 462; Conservative 0; Mismatches 5; Indels 1; Gaps 1;			
OY	838	AGATTCAAATTCCTGAGAGCTGCCCTTACCAGAAGCTCGACGCTTGTTCAGAAGACCT	897
Dd	467	AGATTCAAATTCCTGAGAGCTGCCCTTACCAGAAGCTCGACGCTTGTTCAGAAGACCT	408
OY	898	AATGCACATTGGATTTATTCCTCGAGCTTAATAAAGGAGCTGACTGATGAGAGGCTCTATA	957
Dd	407	AATGCACATTGGATTTATTCCTCGAGCTTAATAAAGGAGCTGACTGATGAGAGGCTCTATA	348
OY	958	AGTTTCTGTGAAGAAGATATGTTTCAACGAMATGGGGTGAACAAGGCGCATGAGAGAG	1017
Dd	347	AGTTTCTGTGAAGAAGATATGTTTCAATGAMATGGGGTGAACAAGGCGCATGAGAGAG	288
OY	1018	ATCAATCTGCCAAGATTAATTCCTCCGCAAGAAAGACTCGAGTCTTTTCAAGCCAACT	1077
Dd	287	ATCAATCTGCCAAGATTAATTCCTCCGCAAGAAAGACTCGAGTCTTTTCAAGCCAACT	228
OY	1078	GCCACACATCATGACGCGCTAAAACGGAAGAGAGCTTCGATATAAACAGCAAGGCGCT	1137
Dd	227	GCCACACATCATGACGCGCTAAAACGGAAGAGAGCTTCGATATAAACAGCAAGGCGCT	168

QY	1138	GCACAAGAAACAAACGCTGTGGAAACAAATATCTGGATGCTGTGATGACAA	1197
Db	167	GCACAAGAAACAAACGCTGTGGAAACAAATATCTGGATGCTGTGATGACAA	108
QY	1198	CTACGACTACGAAGCAGCGGTGATGATCTGCTAGATTATTAACCTCCGTGT	1257
Db	107	CTACGACTACGAAGCAGCGGTGATGATCTGCTAGATTATTAACCTCCGTGT	49
QY	1258	TTTACTCAGAGCTTTGGTAAAGTGGCCATCTTCACGCTGGGTA	1305
Db	48	TTTACTCAGAGCTTTGGTAAAGTGGCCATCTTCACGCTGGGTA	1
RESULT 11			
LOCUS	AM774700	733 bp	mRNA
DEFINITION	EST333851 KV3 Medicago truncatula cDNA clone pKV3-23121, mRNA sequence.		
ACCESSION	AM774700		
VERSION	AM774700.1		
KEYWORDS	GI:7718617		
SOURCE	EST.		
ORGANISM	barrel medic.		
	Medicago truncatula		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.		
REFERENCE	1 (bases 1 to 733)		
AUTHORS	Vandenbosch,K., Hurt,J., Moore,J., Beremand,P., Peng,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and Fraser,C.M.		
TITLE	ESTs from roots of Medicago truncatula after Rhizobium inoculation		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Vandenbosch K Department of Biology Texas A&M University College Station, TX 77843-3258, USA Tel: 409 845 7707 Fax: 409 845 2891 Email: kate@mail.bio.tamu.edu Texas A&M EST name:T258119e Tick sequence name:MT5BE719K More information is available at: http://chryslie.tamu.edu/medicago Seq primer: SKmod (CTA GAA CTA gta gat CC).		
FEATURES			
Source	1..733		
	/organism="Medicago truncatula"		
	/cultivar="genotype A17"		
	/db_xref="taxon:3880"		
	/clone_lib="pKV3-23121"		
	/clone_id="KV3"		
	/tissue_type="Seedling roots"		
	/dev_stage="3 days post-inoculation with Sinorhizobium meliloti"		
	/lab_host="E. coli strain XL0R"		
	/note="Vector: pBluescript SK-; Site:1: EcoRI; Site:2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stragene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-zap phage using Ex-assist helper phage and propagated in XL0R cells."		
BASE COUNT	222 a	125 c	183 g
ORIGIN	203 t		
Query Match	30.4%	Score 420.4	DB: 9; Length 733;
Best Local Similarity	75.4%	Pred. No. 2.8e-79;	
Matches 523:	Conservative 0;	Mismatches 171;	Indels 0;
	Gaps 0;		
36	GATGGGCAATCAAGGCTTTGACGAAACTGCTGGCGACATCGCCCAAGCGATGAAGCA	95	

```

Db 40 GATGGGTATTAGAGGTTTACGAAGCTTTAGCTGATTAATGCTCCCAATGATGAAGA 99
QY 96 GCAGAAATTGAGAGCTACTTGGCCCAAAATGCGCTGCACCCAGCATGACATATA 155
Db 100 GAACAAATTCGAATCTTACTTTGGGCTTAAGATGCTGTTGATGATGATGACATTTA 159
QY 156 CCAGTTCTGATTTAGTTGGAGGAGGATGGAACCTCTCAAAATGAAGTGGTGA 215
Db 160 CAGATTCTTATTGTTGGGGAGAGATGGAACCTGGAATGTTGACTAATGAAGCTGCTGA 219
QY 216 ACTCACTAGTCATTTGCAAGGAATGTTCAACCGACAAATTAAGATTACTGAAGCGGGAT 275
Db 220 ACTACTGTCATTTGCAAGGAATGTTGGCGGACAAATCAGACTCTGAAGCGGGAT 279
QY 276 CAAGCCAGTTTATGTTTGTGATGCAAGCTTCCTGATATGAAGAAACAGAGCTTGCTAA 335
Db 280 GAAGCCAGATATGTTTGTGATGCAAGCTTCCTGATATGAAGAAATCAAGAGCTGAAA 339
QY 336 AAGATACATCAAAAGAGATGATGCAACCAAGATCTGACTGAGCAGAGAGTGAAGA 395
Db 340 AGCTCTCTCAAGAGAGCTGAGGCTACCGCGGTTTGACAGAGAGCTTAGAGGCTGACAA 399
QY 396 TAAAGATGCGATTGAAAAATTGAGCAGAGAGCTGTAAAGGTCAACAGGCAACACGA 455
Db 400 TAAGGAAGATTTGAAAAATTCAGTAAACGACAGTGAAGGTGACAAAGCAATTAATGA 459
QY 456 AAGTTGTAAAGCGCTATTAGACTTATGGGGTTCCTGTTGATGAGGACCTCTTGAGC 515
Db 460 CCAGTGCACAAAGACCTTTGAGACTGAGTGCCTGCTGTGAGGACCCCTCAAGAGC 519
QY 516 AGAAGCAGATGTCAGCCCTTGTGATTAACGATTAAGGTGCTGCTGTTGCTTGAAGA 575
Db 520 AGAGGCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGA 579
QY 576 TATGACTGCTTACTTTTGGGGCTCCAGGCTTCTGCTGCTTAAATGATGCTCAAGTTC 635
Db 580 CATGATCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 639
QY 636 CAGAAATATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 695
Db 640 AAGAAAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 699
QY 696 CATGACACATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
Db 700 CCTGACACATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733

RESULT 12
Bg263233 539 bp mRNA linear EST 16-FEB-2001
LOCUS WHE2339.F04.L072S Wheat pre-anthesis spike cDNA library Trilicium
DEFINITION aestivum cDNA clone WHE2339_F04.L07, mRNA sequence.
ACCESSION Bg263233
VERSION Bg263233.1 GI:12865185
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Trilicium aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
; Triticeae; Triticum.
1 (bases 1 to 539)
Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes - Pre-anthesis spike cDNA library
Unpublished (2000)
CONTACT: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 510595773
Fax: 510595818

```

```

FEATURES
source
1..539
/organism="Trilicium aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE2339.F04.L07"
/clone_lib="Wheat pre-anthesis spike cDNA library"
/tissue_type="Spike before anthesis"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/note="Vector: lambda uni-zap XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
greenhouse. Whole spike with awns trimmed, white, green
and yellow anther were collected and total RNA, and
poly(A) RNA were prepared, a cDNA library was made, and
the cDNA clones were in vivo excised to give pluscript
phagemids in the T7 Close lab (Choi, Close, Fenton) at
the University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
BASE COUNT 163 a 108 c 141 g 126 t 1 others
ORIGIN
Query Match: 29.8%; Score 411.6; DB 10; Length 539;
Best Local Similarity 85.2%; Pred. No. 2,1e-77;
Matches 459; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 112 TACTTCGCGCGCAAAATCGCCGTGAGCGCCAGATGATATACGATTTCTGATTTGA 171
Db 1 TACTTCGCGCGCGCAATCGCCGTGAGCGCCAGATGATATACGATTTCTGATTTGA 60
QY 172 GTTGAAGGACGAGCATGCAATCTTCAAAATGAAGCTGAGAGCTAGTCAATTGG 231
Db 61 GTTGAAGGACGAGCATGCAATCTTCAAAATGAAGCTGAGAGCTAGTCAATTGG 120
QY 232 CAAGAAATGTTCAACCGGACATTAAGTCTGAAACGGGAATCAAGCCGTTTATGT 291
Db 121 CAAGAAATGTTCAACCGGACATTAAGTCTGAAACGGGAATTAACGATATGT 180
QY 292 TTTGATGCGAAGCTCTCTGATATGAGAAACAGAGCTTCTTAAAGATCTCAAAAGA 351
Db 181 TTTGATGCGAAGCTCTCTGATATGAGAAACAGAGCTTCTTAAAGATCTCAAAAGA 240
QY 352 GATGATGCAACCAAGATGACATGAGCAGTGAAGCTGAGATGAAGATGAAGATGA 411
Db 241 AATGAGAGCAAGAGAGCTGTAAGAGTCAAGAGCAACAGAGATGTAAGAGCTGA 300
QY 412 AATGAGAGCAAGAGAGCTGTAAGAGTCAAGAGCAACAGAGATGTAAGAGCTGA 471
Db 301 AATGAGAGCAAGAGAGCTGTAAGAGTCAAGAGCAACAGAGATGTAAGAGCTGA 360
QY 472 TTAAGACTTATGCGGCTTCTGATGAGAGCACTTCTGAGAGAGAGAGATGGA 531
Db 361 CTAAAGACTGAGAGGCTTCTGATGAGAGCACTTCTGAGAGAGAGAGATGGA 420
QY 532 GCCCTTTCATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
Db 421 GCCCTTTCATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 592 TTTGGGGCTCCACGGTCTTCTGATATGATGATGATGATGATGATGATGATGATGAT 650
Db 481 TTTGGGGCTCCACGGTCTTCTGATATGATGATGATGATGATGATGATGATGATGAT 539

RESULT 13
AW562517/c 475 bp mRNA linear EST 10-MAR-2000
LOCUS AW562517
DEFINITION 66065H06.X2 660 - Mixed stages of anther and pollen Zea mays cDNA,
mRNA sequence.

```



```

ACCESSION      AM562517
VERSION        AM562517.1  GI:7216395
KEYWORDS       EST.
SOURCE         Zea mays.
ORGANISM       Zea mays.
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
                clade; Panicoidae; Andropogoneae; Zea.
REFERENCE      1 (bases 1 to 475)
AUTHORS        Walbot,V.
TITLE          Maize ESTs from various cDNA libraries sequenced at Stanford
                University
JOURNAL        Unpublished (1999)
COMMENT        Contact: Walbot V
                Department of Biological Sciences
                Stanford University
                855 California Ave, Palo Alto, CA 94304, USA
                Tel: 650 723 2227
                Fax: 650 725 8221
                Email: walbot@stanford.edu
                Plate: 660065 row: H column: 06.
FEATURES
  source
    1..475
    /organism="Zea mays"
    /cultivar="Ohio43"
    /db_xref="taxon:4577"
    /clone_lib="660 - Mixed stages of anther and pollen"
    /tissue_type="whole premeiotic anthers to pollen shed"
    /dev_stage="premeiotic anthers to pollen shed"
    /lab_host="XLOLR"
    /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
    Site_2: XhoI; Anther and pollen cDNA library.
    Directionally sequenced with 5' end at the EcoRI site.
    Created by Amie Franklin."
BASE COUNT      124 a 110 c 94 g 147 t
ORIGIN
Query Match      29.7%; Score 409.8; DB 9; Length 475;
Best Local Similarity 98.2%; Pred. No. 5.1e-77;
Matches 425; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 931 ACAGCAACCTGATGAGGAGGCTGCATTAAGTTCCTGCTGTAAGAATATGCTTCACAGCA 990
    |||||||
DB 475 ACTGCACCTGATGAGGAGGCTGCATTAAGTTCCTGCTGTAAGAATATGCTTCATGAA 416
QY 991 GATCGGGTGCAAAAGSCCATAGAGAAGATCAATCTGCCAAGATTAATCTGTCGCAAGA 1050
    |||||||
DB 415 GATCGGGTGCAAAAGSCCATAGAGAAGATCAATCTGCCAAGATTAATCTGTCGCAAGA 356
QY 1051 AGACTGAGTCCCTTTTCAAGCAACTGCCACCATCGCACCGCTTAAGCGGAAGAG 1110
    |||||||
DB 355 AGACTGAGTCCCTTTTCAAGCAACTGCCACCATCGCACCGCTTAAGCGGAAGAG 296
QY 1111 ACTTCGATTAACAACAAGCAGGAGCTGCGAACAAGAAAACAAGGCTGTGGAAGAG 1170
    |||||||
DB 295 ACTTCGATTAACAACAAGCAGGAGCTGCGAACAAGAAAACAAGGCTGTGGAAGAG 236
QY 1171 AATAATCTTGATGCTGTGATGTACACTACACATACGAAGACAGCGTGCCTATAC 1230
    |||||||
DB 235 AATAATCTTGATGCTGTGATGTACACTACACATACGAAGACAGCGTGCCTATAC 176
QY 1231 TTGCGTAGATTAATTAACCTCCCTGTTTAACCTACAGAGTTTGTAAGAAGTTCGCCCATG 1290
    |||||||
DB 175 TTGCGTAGATTAATTAACCTCCCTGTTTAACCTACAGAGTTTGTAAGAAGTTCGCCCATG 116
QY 1291 TTTCACGCTGGGTAAGTGTGTGTTGAAGAGCTTGTTACCAAGTAAACAACCTT 1350
    |||||||
DB 115 TTTCACGCTGGGTAAGTGTGTGTTGAAGAGCTTGTTACCAAGTAAACAACCTT 57
QY 1351 ATGCTGTGTTT 1363
    |||||||
DB 56 ATGCTGTGTTT 44

```

```

RESULT 14
LOCUS      AM288784
DEFINITION 70710LOC02.x5 707 - Mixed adult tissues from Walbot lab (SK) Zea
ACCESSION  AM288784
VERSION    AM288784.1  GI:6695706
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Zea mays.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoidae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 414)
AUTHORS    Walbot,V.
TITLE      Maize ESTs from various cDNA libraries sequenced at Stanford
            University
JOURNAL    Unpublished (1999)
COMMENT    Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 70710 row: C column: 02.
FEATURES
  source
    1..414
    /organism="Zea mays"
    /cultivar="W23"
    /db_xref="taxon:4577"
    /clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
    "
    /tissue_type="tassel, kernel, silk, husk, root, leaf"
    /dev_stage="adult"
    /lab_host="DH10B"
    /note="Organ: tassel, kernel, silk, husk, root, leaf;
    Vector: pGAD10; Site_1: EcoRI, cDNA library from fully
    differentiated maize tissues from an active Mutator
    plant. Tissue ratio is 4/21/1/1/1 (tassel, kernel, silk,
    husk, root, leaf). Unidirectionally cloned."
BASE COUNT      120 a 82 c 101 g 109 t 2 others
ORIGIN
Query Match      29.1%; Score 402.2; DB 9; Length 414;
Best Local Similarity 98.8%; Pred. No. 2.1e-75;
Matches 404; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 601 CCACGGTTCCTTCGTCATTTATGATCCAAAGTTCAGAAATTAACCTGATGGAATT 660
    |||||||
DB 6 CCACGGTTCCTTCGTCATTTATGATCCAAAGTTCAGAAATTAACCTGATGGAATT 65
QY 661 GATGTGCCAAGGTTTGGAGAGGCTTGAACATCACATGACCAAGTTCATGATTTGTC 720
    |||||||
DB 66 GATGTGCCAAGGTTTGGAGAGGCTTGAACATCACATGACCAAGTTCATGATTTGTC 125
QY 721 ATCTGTGTGATGTGACTATTGTGATAGCATCAAGATATCGGGGGCAAAACGCTCTG 780
    |||||||
DB 126 ATCTGTGTGATGTGACTATTGTGATAGCATCAAGATATCGGGGGCAAAACGCTCTG 185
QY 781 AACTATTTCGTCACATGGTGTCCATAGAAAGCATCTGGAGAAATCTTAATAAAGACAGA 840
    |||||||
DB 186 AACTATTTCGTCACATGGTGTCCATAGAAAGCATCTGGAGAAATCTTAATAAAGACAGA 245
QY 841 TATCAATTCCTGAGAGCTGCGCTTACCAAGAGCTGACGCTGTTCACAGACCTAAT 900
    |||||||
DB 246 TACCAATTCCTGAGAGCTGCGCTTACCAAGAGCTGACGCTGTTCACAGACCTAAT 305
QY 901 GTCAATTTGATATTCCTAGCTAAATATGAGTACCTGATGAGAGAGGCTCATAGT 960
    |||||||
DB 306 GTCAATTTGATATTCCTAGCTAAATATGAGTACCTGATGAGAGAGGCTCATAGT 365

```

QY      961 TTCCTGTTAAAGATTAATGGTTTCAACGAAGATCGGTGACCAAGGCCA 1009  
|||||  
|||  
Db      366 TTCCGTGTTAAAGATTAATGGTTTCAACGAAGATCGGCTGACCAAGGCCA 414

RESULT	15
A1065689/c	
LOCUS	A1065689 456 bp mRNA linear EST 24-JUL-1998
DEFINITION	ag1ff12.x1 maize inflorescence immature ear library Zea mays cDNA
ACCESSION	A1065689
VERSION	A1065689.1 GI:3341096
KEYWORDS	EST.
SOURCE	Zea mays.
ORGANISM	Zea mays

ONLINE.COM

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

Embryophyta: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:  
Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: PACC  
clade: Pennicliaceae: Andropogoneae: Zea.  
1 (bases 1 to 456)  
Schutz, K., de la Bastide, M., Gnoj, J., Hebermann, K., Huang, E. N.,  
Pantell, L. D., Dedina, N., Mattiensen, R. and McCombie, W. R.  
Expressed sequence tags from *Z. mays*  
Unpublished (1998)  
Contact: W. Richard McCombie

CONTACT: W. RICHARD MCCOMB  
Lila Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel.: 516 367 8884  
Fax: 516 367 8874  
Email: [mccomb@cspl.org](mailto:mccomb@cspl.org)  
Plate: a91 row: f column: 12  
Seq primer: M13 forward universal -21  
High quality sequence: stop: 456.

FEATURES	Location/Qualifiers
source	1. .456

```

SOURCE
1: /organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ag11i12"
/clone.lib="maize inflorescence immature ear library"
/sex="female"
/tissue_type="immature ear"
/note="Vector: pRINTEGRATE SK+ (X52325): Site.1: XhoI;
Site.2: EcoRI; This library is described in Schmidt, Hake,
et al., (1993) Plant Cell 5:729-737. cDNAs are
directionally cloned into the XhoI and EcoRI sites; XhoI
is near the polyA tail. Most reads from this library are
3' in direction. Additional information on this library as
well as ftp access to all sequences can be found at
http://www.cshl.org/maizegenome"
BASE COUNT
112 a 104 c 86 g 152 t 2 others
ORIGIN

```

Query Match	28.7%	Score 397	DB 9	Length 456
Similarity	91.9%	Pred. No. 2.7e-74		
Matches 418	Conservative	0	Mismatches 37	Indels 0
				Gaps 0

QY	909	GGATATTCCTCAGCAAAATAAGCATGCACCTGATGAGAGGGGTCATCAATGATTTCTGT	968
Db	456	GGATATTCCTGAGCTAAACTGACCTGCACCTGATGAGAGGGTCTATTAATGTTCTGT	397
QY	969	AAAAGATAATGTTTCAACGAGATCGGATGACAAAGGCCATAGAGAAATCAAAATGTC	1028
Db	396	AAAAGATAATGTTTCAACGAGAAATCGGGGACAAAGGCCATAGAAATCAATCTGC	337
QY	1029	CAAGATAAATTCGTCGCAAGSAGACTGAGTCTTTTTCAGCCAACTGCCACACATC	1088
Db	336	CAAGATAAATTCGTCGCAAGSAGACTGAGTCTTTTTCAGCCAACTGCCACACATC	277
QY	1089	AGCAGCGCAAAAGCGAAGSAGACTGAGTAAATAACCAAGCAAGCGCTGCGACAAACA	1148
Db	276	AGCAGCGCTAAACGGAAGSAGACTTGGATTAATAACCAAGCAAGCGCTGCGACAAACA	217

QY	1149	AACAAAGCGTGGTGGAAAGAAATAATCTGGAGGCTGGATGCTACACTACGCATACG	1208
Db	216	AACAAGAGCGCTGGTGGAAAGAAATATATCTTGGATCTGTTAGTACACTACGCATACG	157
QY	1209	AAMGAGCGGCGGCGGTGATCACTTCGCTTGAGATTTATTTAACTCCCTGTTTAACTCAGAG	1266
Db	156	AAMGAGCGGCGGCGGTGATCACTTCGCTTGAGATTTATTTAACTCCCTGTTTAACTCAGT	97
QY	1265	CTTTGCTAAAGTTCGCCCATGTTTCAAGCGGGGTAGTACTTCTGTTGAGAGATT	1328
Db	96	TGCAAGAGATNGGTGATCCACAGTAAACAACACTATCGCTGTTTTTTATGTTTGAAGAGATN	37
QY	1329	GGTGTACCAAGTAACAAACACTTATCGCTGTTTTTT	1363
Db	36	GGTGTACCAAGTAACAAACACTTATCGCTGTTTTTT	2

Search completed: November 5, 2002, 13:36:58  
Job time : 1195.61 secs





319 TTACTGGAAGCGGGAATCAAGCCACTTATGTTTTGATGGCAAGCCCTCGATATGGAAG 378  
Db 307 TTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTGATGGCAAGCCCTCGATATGGAAG 366  
379 AAACAGAACTTGGCTAAAGATPACTCAAAAGAGATGATGCACCAAGATCTGACTGAG 438  
Db 367 AAACAGAACTTGGCTAAAGATPACTCAAAAGAGATGATGCACCAAGATCTGACTGAG 426  
439 GCAGTAGAGTAGAGATGAAGATCGATTTGAAAATTTGACCAAGAGAGACTGTAAAGGTC 498  
Db 427 GCAGTAGAGTAGAGATGAAGATCGATTTGAAAATTTGACCAAGAGAGACTGTAAAGGTC 486  
499 ACAAGGCAACCAAGGAAGATTTAAGGACTATTAAAGCTTTAGGGGCTTCTGTTGTA 558  
Db 487 ACAAGGCAACCAAGGAAGATTTAAGGACTATTAAAGCTTTAGGGGCTTCTGTTGTA 546  
559 GAGGCACTTCTGGAAGCGAAGATGTGACGCCCTTTGCATTAAGGATTAAGGTGTC 618  
Db 547 GAGGCACTTCTGGAAGCGAAGATGTGACGCCCTTTGCATTAAGGATTAAGGTGTC 606  
619 GCTGTGCTTCAAGAGATTAAGGACTCCCTTACTTTTGGGGCTCCACGGCTTCTGTCAT 678  
Db 607 GCTGTGCTTCAAGAGATTAAGGACTCCCTTACTTTTGGGGCTCCACGGCTTCTGTCAT 666  
679 TTATGTGATCCAAAGTTCACAGAAAATACCTGTGATGCAATTTGATGTTGCCAAGGTTTG 738  
Db 667 TTATGTGATCCAAAGTTCACAGAAAATACCTGTGATGCAATTTGATGTTGCCAAGGTTTG 726  
739 GAGGAGCTTGAACCTCACCATGGACCAGTTTCATTTGATTTGTCATTCCTGTGGATGTAC 798  
Db 727 GAGGAGCTTGAACCTCACCATGGACCAGTTTCATTTGATTTGTCATTCCTGTGGATGTAC 786  
799 TATTGTATAGCATCAAAAGGTATCGGGGGCAACAGCTGGAACATTTATTCGTCAACAT 858  
Db 787 TATTGTATAGCATCAAAAGGTATCGGGGGCAACAGCTGGAACATTTATTCGTCAACAT 846  
859 GGGTCCCATGAAGACATCTTGGAGATCTTAAATAAGACATATCAAAATTCCTGAGAC 918  
Db 847 GGGTCCCATGAAGACATCTTGGAGATCTTAAATAAGACATATCAAAATTCCTGAGAC 906  
919 TGGCCTTACCAAGAGAGTGCAGCGTGTGTTCAAGAGAGCTTATGTCACCTTGGATATTCCT 978  
Db 907 TGGCCTTACCAAGAGAGTGCAGCGTGTGTTCAAGAGAGCTTATGTCACCTTGGATATTCCT 966  
979 GAGCTAAATGAGTGCAGCTGATGAGAGAGGCTCATTAAGTTCTGCTGTAAGATTAAT 1038  
Db 967 GAGCTAAATGAGTGCAGCTGATGAGAGAGGCTCATTAAGTTCTGCTGTAAGATTAAT 1026  
1039 GGTTCATGAGAGATCGGGTGCACAAAGGCCATAGAGAAAGATCAAAATCTGCCAAGATTA 1098  
Db 1027 GGTTCATGAGAGATCGGGTGCACAAAGGCCATAGAGAAAGATCAAAATCTGCCAAGATTA 1086  
1099 TCGTCGCAAGAGAGATCGAGTCTTTTTCNAAGCAACTGCGCACCATCGACCGGTA 1158  
Db 1087 TCGTCGCAAGAGAGATCGAGTCTTTTTCNAAGCAACTGCGCACCATCGACCGGTA 1146  
1159 AAACGGAAGAGACTTGGATTAACAAGCAAGGCACTGCGAACAAGAAAAGAGCT 1218  
Db 1147 AAACGGAAGAGACTTGGATTAACAAGCAAGGCACTGCGAACAAGAAAAGAGCT 1206  
1219 GGTGGAAGAGAGAAATATCTTGGATGCTGATGATACAACTACGACTACGAAACACAGCG 1278  
Db 1207 GGTGGAAGAGAGAAATATCTTGGATGCTGATGATACAACTACGACTACGAAACACAGCG 1266  
1279 TGGCATGATCACTTGGCTAGATTAATTAATCTCCCTGTTTAACTCGAGACTTGGTGAA 1338  
Db 1267 TGGCATGATCACTTGGCTAGATTAATTAATCTCCCTGTTTAACTCGAGACTTGGTGAA 1326  
1339 AGTTTGGCCATGTTTCAAGCTGGGGTAAGTTAGTTGTTTGAAGAGATTGGTGTACAA 1398  
Db 1327 AGTTTGGCCATGTTTCAAGCTGGGGTAAGTTAGTTGTTTGAAGAGATTGGTGTACAA 1386  
1399 GTAACAACACTATCGCTGTTTTTACTTCTGTCTTGAAGATGATATGSCCGTAAAA 1458

Db 1387 GTAACAACACTATCGCTGTTTTTACTTCTGTCTTGAAGTAAAAA 1446  
QY 1459 AAAAAAAAAAAAAAAAAA 1475  
Db 1447 AAAAAAAAAAAAAAAAAA 1463  
RESULT 3  
AR152404  
LOCUS AR152404 1541 bp DNA linear PAT 08-Aug-2001  
DEFINITION Sequence 3 from patent US 6232527.  
ACCESSION AR152404  
VERSION AR152404.1 GI:15118454  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1541)  
AUTHORS Mahajan,P.B.  
TITLE Maize Rad2/REN-1 orthologues and uses thereof  
JOURNAL Patent: US 6232527-A 3 15-May-2001;  
FEATURES  
source  
1..1541  
/organism="unknown"  
BASE COUNT 473 a 308 c 377 g 383 t  
ORIGIN  
Query Match 94.8%; Score 1401.2; DB 6; Length 1541;  
Best local Similarity 99.1%; Pred. No. 0;  
Matches 1409; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 33 GCGGTTTTCGGCCACACCGCGCTCAGCCGCCGCCGCCGACACCCGCCACACGCCGCCGCGAGA 92  
Db 15 GCGGTTTTCGGCCACACCGCGCTCAGCCGCCGCCGCCGACACCCGCCACACGCCGCCGCGAGA 74  
QY 93 CGAGATGGGCATCAAGGTTTGAGCAAACTGCTGGCGGACATGGCGCCCAAGGCGATGAA 152  
Db 75 CGAGATGGGCATCAAGGTTTGAGCAAACTGCTGGCGGACATGGCGCCCAAGGCGATGAA 134  
QY 153 GGAGCAGAAGTTTGCAGAGCTTCTTGGCGCCGCAAAATCGCCCTGCAGCGCAGATGAGCAT 212  
Db 135 GGAGCAGAAGTTTGCAGAGCTTCTTGGCGCCGCAAAATCGCCCTGCAGCGCAGATGAGCAT 194  
QY 213 CTACAGTTCTGTAGTAGTGTGGAAGGACAGGCAATGGAACCTCCACAATGGAAGCTGG 272  
Db 195 ATACCAAGTTCTGTAGTTAGTTGGAAGGACAGGCAATGGAACCTCCACAATGGAAGCTGG 254  
QY 273 TGAAGTCACTAGTCAATTTGCAAGGATGTTCAACCGGCAATAAGATTAAGTGAAGCGGG 332  
Db 255 TGAAGTCACTAGTCAATTTGCAAGGATGTTCAACCGGCAATAAGATTAAGTGAAGCGGG 314  
QY 333 AATCAAGCAGTTTATGTTTGTGATGGCAAGCTCTGTATATGAAGAAACAAGACTTGC 392  
Db 315 AATCAAGCAGTTTATGTTTGTGATGGCAAGCTCTGTATATGAAGAAACAAGACTTGC 374  
QY 393 TAAAGATACCTCAAAAAGAGATGATGCAACCAAGATGACTGCTGAGGCAAGAGGTAGG 452  
Db 375 TAAAGATACCTCAAAAAGAGATGATGCAACCAAGATGACTGCTGAGGCAAGAGGTAGG 434  
QY 453 AGATTAAGATCGATTTGAAAATTTGAGCAAGAGAGACTGTAAGGTCAACAAGCAACAA 512  
Db 435 AGATTAAGATCGATTTGAAAATTTGAGCAAGAGAGACTGTAAGGTCAACAAGCAACAA 494  
QY 513 CGAAGATTGTAAGGACTATTAAAGCTTAAGGGGTTCTTGTATAGAGCACTTCTGA 572  
Db 495 CGAAGATTGTAAGGAGGCTATTAAAGCTTAAGGGGTTCTTGTATAGAGCACTTCTGA 554  
QY 573 AGCAGAAGCAGAATGTGAGCCCTTTGCAATAAAGATTAAGGTGTTGCTGTTGCTTGA 632  
Db 555 AGCAGAAGCAGAATGTGAGCCCTTTGCAATTAAGGTGTTGCTGTTGCTTGA 614  
QY 633 AGATTAAGACTCCTTACTTTTGGGGCTCCACGGTCTCTTCGTCATTTAATGATCCAA 692

Db	615	AGATATGGA	CTCCCTTACTTTTGGGGGCTCCAGGTTCTTGCTATTAATGATCCAA	674
OY	693	TTCCAGAA	AAATTAACCTGTGATGGAATTTGATGTTGCCAAGCTTTTGGAGAGCTTGA	752
Db	675	TTCCAGAA	AAATTAACCTGTGATGGAATTTGATGTTGCCAAGCTTTTGGAGAGCTTGA	734
OY	753	CACCATGA	CAACGATTCATGATTTGTGATCCTGTGTGATGATGATCTATTTGATAGCAT	812
Db	735	CACCATGA	CAACGATTCATGATTTGTGATCCTGTGTGATGATGATCTATTTGATAGCAT	794
OY	813	CAAAAGTA	TGCGGGGGGCAAGAGCTCTGAAACTTATTTCTGCAACATGSGCTCCAT	872
Db	795	CAAAAGTA	TGCGGGGGGCAAGAGCTCTGAAACTTATTTCTGCAACATGSGCTCCAT	854
OY	873	CATCTGGA	GAGATCTTATTAAGACAGATATCAATTCCTGTGAGCATGSGCTTACCA	932
Db	855	CATCTGGA	GAGATCTTATTAAGACAGATATCAATTCCTGTGAGCATGSGCTTACCA	914
OY	933	AGCTGAGC	CTTGTTCAGAGAGCCTTAATGTCACTTTGGATATTTCTGTGAGCTAAAA	992
Db	915	AGCTGAGC	CTTGTTCAGAGAGCCTTAATGTCACTTTGGATATTTCTGTGAGCTAAAA	974
OY	993	TGCACTTA	TATGAGGAGGCTTCATAGTTTCTGTGTAAGATTAATGCTTTCAATGAGA	1052
Db	975	TGCACTTA	TATGAGGAGGCTTCATAGTTTCTGTGTAAGATTAATGCTTTCAATGAGA	1034
OY	1053	TCGGGTGA	CAAAAGGCGCATAGAGAGATCAAAATCTGCCAAGATTAATGCTGCCA	1112
Db	1035	TCGGGTGA	CAAAAGGCGCATAGAGAGATCAAAATCTGCCAAGATTAATGCTGCCA	1094
OY	1113	ACTCGAGT	CTTTTTCAGAGCCACTGCGCACATCAGCAGCGCTTAAGAAAGGAGAGAC	1172
Db	1095	ACTCGAGT	CTTTTTCAGAGCCACTGCGCACATCAGCAGCGCTTAAGAAAGGAGAGAC	1154
OY	1173	TTTCGATA	AAAAACAACGACGAGCTGGCAACAAGAAAAACAAGGCTGGTGGCAAGAA	1232
Db	1155	TTTCGATA	AAAAACAACGACGAGCTGGCAACAAGAAAAACAAGGCTGGTGGCAAGAA	1214
OY	1233	ATAATCTT	GGATGCTTGTATGTAACAATACTACGACTACGAAGACAGCGGTGGCAT	1292
Db	1215	ATAATCTT	GGATGCTTGTATGTAACAATACTACGACTACGAAGACAGCGGTGGCAT	1274
OY	1293	CGCCGAGA	TTATTTAATACCTCCGTTTATACACAGACCTTTGGTGAAGTTTCCAT	1352
Db	1275	CGCCGAGA	TTATTTAATACCTCCGTTTATACACAGACCTTTGGTGAAGTTTCCAT	1334
OY	1353	TCAAACCT	GGCGGTAAGTGTGCTTTGGAAGATTTGCTACCAAGTAACAAACTTAT	1412
Db	1335	TCAAACCT	GGCGGTAAGTGTGCTTTGGAAGATTTGCTACCAAGTAACAAACTTAT	1394
OY	1413	CGCTGTTT	TTTACTTCTGTCTTTGAAGTATGATGCGAGT 1454	
Db	1395	CGCTGTTT	TTTACTTCTGTCTTTGAAGTATGATGCGAGT 1436	
RESULT 4				
LOCUS	ARI52405		1381 bp	DNA linear PAT 08-AUG-2001
DEFINITION	Sequence 5 from patent US 6232527.			
ACCESSION	ARI52405			
VERSION	ARI52405.1		GI:15118455	
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 1381)			
AUTHORS	Mahajan, P. B.			
TITLE	Maize Rad2/FEN-1 orthologues and uses thereof			
JOURNAL	Patent: US 6232527-A 5 15-MAY-2001;			
FEATURES	Location/Qualifiers			
	1..1381			
	/organism="unknown"			

BASE COUNT	441 a	269 c	346 g	325 t	ORIGIN
Query Match	90.3%	Score 1334.2	DB 6	Length 1381	
Best Local Similarity	98.7%	Pred. No. 0			
Matches 1345	Conservative	0	Mismatches 18	Indels 0	Gaps
Qy	61	CGCGCCCGCCACCCCGCCACAGCCGCCGAGAGAGCAATGGGCGATCAAGGGTTTGACGAA	120		
Db	1	CGACCCACACGCTCCGGCCACAGCCGCCGAGAGAGATGGGCGATCAAGGGTTTGACGAA	60		
Qy	121	CTGCTGGGGACAATGCGCCCAAGCGGATGTAAGAGCAGAACTTGGAGACTCTCGGC	180		
Db	61	CTGCTGGGGACAATGCGCCCAAGGGGATGTAAGAGCAGAACTTGGAGACTCTCGGC	120		
Qy	181	CGCAAAATCGCCGTCGACGCGCACATGAGATCTACCACTTCTCATATGATGGTGAAG	240		
Db	121	CGCAAAATCGCCGTCGACGCGCACATGAGATCTACCACTTCTCATATGATGGTGAAG	180		
Qy	241	ACAGGCATGGAACCTCTCACAAATGAAGCTGTGAAGTCACTAGTCAATTTGCAAGGAATG	300		
Db	181	ACAGGCATGGAACCTCTCACAAATGAAGCTGTGAAGTCACTAGTCAATTTGCAAGGAATG	240		
Qy	301	TTTCAACCGGCAATTAACATTAAGTGAAGCGGGCAATCAACCCAGTTATGTTTGAATGC	360		
Db	241	TTTCAACCGGCAATTAACATTAAGTGAAGCGGGCAATCAACCCAGTTATGTTTGAATGC	300		
Qy	361	AAGCCTCTGTATATGAAGAAACAAGAACTTGCTAAAGATACTCAAAAAGAGATGATCA	420		
Db	301	AAGCCTCTGTATATGAAGAAACAAGAACTTGCTAAAGATACTCAAAAAGAGATGATCA	360		
Qy	421	ACCAAAAGTCTGACTGAGCGAGTAGAGATGAGATGAAGATGGCATTTGAAAATTTGAGC	480		
Db	361	ACCAAAAGTCTGACTGAGCGAGTAGAGATGAGATGAAGATGGCATTTGAAAATTTGAGC	420		
Qy	481	AAGAGACTGTAAAGGTCACACAGCAACACACAGATTTGTAACGACTATTAAAGACTT	540		
Db	421	AAGAGACTGTAAAGGTCACACAGCAACACACAGATTTGTAACGACTATTAAAGACTT	480		
Qy	541	ATGGGGCTTCTGTTGTGTAAGGCACTTCTGAAGCAAGAGCAAGATGTGACGCTTTTGC	600		
Db	481	ATGGGGCTTCTGTTGTGTAAGGCACTTCTGAAGCAAGAGCAAGATGTGACGCTTTTGC	540		
Qy	601	ATAAAGCAATTAAGGTTTGGCGCTTGCTTCAAGATTAAGACACCTTACTTTGGGACT	660		
Db	541	ATAAAGCAATTAAGGTTTGGCGCTTGCTTCAAGATTAAGACACCTTACTTTGGGACT	600		
Qy	661	CCAGCGTCTCTGTCATTTAATGATTCACAGTTCCAGAATAATACCTGTGATGGAATTT	720		
Db	601	CCAGCGTCTCTGTCATTTAATGATTCACAGTTCCAGAATAATACCTGTGATGGAATTT	660		
Qy	721	GATGTGGCAAGGTTTGGAGGAGCTGGAAGTACCATGGACAGTTCATTTATTTGGC	780		
Db	661	GATGTGGCAAGGTTTGGAGGAGCTGGAAGTACCATGGACAGTTCATTTATTTGGC	720		
Qy	781	ATCCGTGTGATGTCATTTGATGATGATCAAGGATATCGGGGGGCAACACGCTGTG	840		
Db	721	ATCCGTGTGATGTCATTTGATGATGATCAAGGATATCGGGGGGCAACACGCTGTG	780		
Qy	841	AAACTTATTCGTCAACATGGBTCCANTAGAAAGCATCTTGSAGATCTTATTAACACAGA	900		
Db	781	AAACTTATTCGTCAACATGGBTCCANTAGAAAGCATCTTGSAGATCTTATTAACACAGA	840		
Qy	901	TATCAAAATTCCTGAGGAGCTGGCCTTACCAGAAGGCTCGAGCGCTGTTCAAGAGGCTTAT	960		
Db	841	TATCAAAATTCCTGAGGAGCTGGCCTTACCAGAAGGCTCGAGCGCTGTTCAAGAGGCTTAT	900		
Qy	961	GTCACCTTTGATATTCCTGAGACTTAAATGAGACTGCACCTGATGAGGAGGCTCTCAATG	1020		
Db	901	GTCACCTTTGATATTCCTGAGACTTAAATGAGACTGCACCTGATGAGGAGGCTCTCAATG	960		
Qy	1021	TTTCTGTGTAAGAAATATGTTTCAATGAAGATTCGGGTGACAAAGCCATTAAGAAATG	1080		

Db 961 TTCTCGTGAATAAGATGTTTCAACGAGATCGGGTGACCAAGGCCATAGAGAATC 1020  
 Oy 1081 AATCTGCCAAGATTAATCTGCGACGAGAGACTCGAGTCTTTTTCAGGCACTGCC 1140  
 Db 1021 AATCTGCCAAGATTAATCTGCGACGAGAGACTCGAGTCTTTTTCAGGCACTGCC 1080  
 Oy 1141 ACCACATCGACACCGCTTAACGGAAGAGACTTGGATTAACACGACGAGCTGCG 1200  
 Db 1081 ACCACATCGACACCGCTTAACGGAAGAGACTTGGATTAACACGACGAGCTGCG 1140  
 Oy 1201 AACCAAAACAAAGGCTGTGGAAGAGAAATATCTTGATCGTGTATACACCTA 1260  
 Db 1141 AACCAAAACAAAGGCTGTGGAAGAGAAATATCTTGATCGTGTATACACCTA 1200  
 Oy 1261 CGACTACGAAACGAGGCTGGCATGATCCTTCGCTAGATTATTAATCCCTGTTTGA 1320  
 Db 1201 CGACTACGAAACGAGGCTGGCATGATCCTTCGCTAGATTATTAATCCCTGTTTGA 1260  
 Oy 1321 ACTCAGACTTTGGTGAAGTTTGGCCATGTTCAAGCTGGGCTAAGTACTGCTTGG 1380  
 Db 1261 ACTCAGACTTTGGTGAAGTTTGGCCATGTTCAAGCTGGGCTAAGTACTGCTTGG 1320  
 Oy 1381 AAGAGATTGTGTACCACTAACAACTTATCGCTGTTT 1423  
 Db 1321 AAGAGATTGTGTACCACTAACAACTTATCGCTGTTT 1363

RESULT 5  
 AB021666 1354 bp mRNA linear PLN 25-MAR-2000  
 LOCUS  
 DEFINITION Oryza sativa OsFEN-1 mRNA for FEN-1, complete cds.  
 ACCESSION AB021666 GI:4587224  
 VERSION FEN-1; endonuclease.  
 KEYWORDS Oryza sativa cDNA to mRNA.  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 AUTHORS Kimura,S., Ueda,T., Hatanaka,M., Takenouchi,M., Hashimoto,J. and Sakaguchi,K.  
 TITLE Plant homologue of flap endonuclease-1: molecular cloning, characterization, and evidence of expression in meristematic tissues  
 JOURNAL Plant Mol. Biol. 42 (3), 415-427 (2000)  
 MEDLINE 20256470  
 AUTHORS Kimura,S., Hashimoto,J. and Sakaguchi,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-DEC-1998) Seisuke Kimura, Science University of Tokyo, Dept. of Applied Biological Science, 2641 Yamazaki, Noda, Chiba 278-8510, Japan (E-mail:j5498703@nodai.ac.jp, Tel:81-471-24-1501(ex.3419), Fax:81-471-23-9767)  
 COMMENT On Apr 17, 1999 this sequence version replaced gi:4062866.

FEATURES  
 source  
 1..1354  
 /organism="Oryza sativa"  
 /db\_xref="taxon:4530"  
 67..1209  
 /gene="OsFEN-1"  
 67..1209  
 /gene="OsFEN-1"  
 /note="OsFEN-1"  
 /note="endonuclease"  
 /codon\_start=1  
 /product="FEN-1"  
 /protein\_id="BAA36171.1"  
 /db\_xref="gi:4587225"  
 /translation="MGIRGLTKLADNAPKAMKQKPEPSYGRRIAYDASMSYIOFLI  
 VYGRGMETLTNEAGEVSHDGMENRIRILEAIKRYVYDPPLKQOELAKRY  
 SKREDATKELTVAVEGKDAIEFKSKVTYKQHNSEKRLTLKGVPAVEAFCEA  
 EAECALCINDVAVASEDMDSLEFAGPRFLRIHMDSSKRLPYMEVEAVKLELE  
 LMDQFDLCILISGCDYDSIKIGIGQATLKLIRHGSIESILENTINRDRQIPEMDP

gene  
 CDS

YOEARLEKPEPNVTLIDIPELKMNAPDEEGLVEFLVKNENFENDRVTKAIEIKFAKNK  
 SSQGRLESEFFKPVYSTVPLKRRKDTSEKPTKAVANKKTKGAGKKK"  
 polyA\_site  
 1354  
 /note="17 A nucleotides"  
 BASE COUNT 415 a 286 c 331 g 322 t  
 ORIGIN

Query Match 62.4%; Score 922.6; DB 8; Length 1354;  
 Best Local Similarity 86.1%; Pred. No. 1,7e-218;  
 Matches 1034; Conservative 0; Mismatches 164; Indels 3; Gaps 1;

Oy 67 CCGGACCCCGCCACAGCCCGCCAGACAGAGATGGGCATCAAGGTTTGAGCAAACTCTG 126  
 Db 37 CCGGACCCCGCCACAGCCCGCCAGACAGAGATGGGCATCAAGGTTTGAGCAAACTCTG 96  
 Oy 127 CGGACAAATGCGCCCAAGGCGATGAAGAGAGCAAGTTTCAGACCTACTTCGCGCCGAAA 186  
 Db 97 CGGACAAATGCGCCCAAGGCGATGAAGAGAGCAAGTTTCAGACCTACTTCGCGCGGCG 156  
 Oy 187 ATCGCCGTCGACGCCAGCATGACATCTACCATGTTCTGATAGTAGTTGGAAGCAGCGC 246  
 Db 157 ATCGCCGTCGACGCCAGCATGACATCTACCATGTTCTGATAGTAGTTGGAAGCAGCGG 216  
 Oy 247 ATGCAAACTCTCACAAATGAAGCTGTGAAGTCACTAGTCACTTTGCAAGAAATTTCAAC 306  
 Db 217 ATGCAAACTCTCACAAAGAGAGAGCGCGGTAAAGTCACTAGTCACTTTGCAAGATGTTCAAC 276  
 Oy 307 CGGACAAATGAATTAATCTGGAAGCGGGAATCAAGCCAGTTTATGTTTATGATGCAAGCCT 366  
 Db 277 CGGACAAATGAATTAATCTGGAAGCGGGAATCAAGCCAGTTTATGTTTATGATGCAAGCCT 336  
 Oy 367 CCGTATGAAGAAACAAAGAACTTGTAAAGATATCTCAAAAGAGATGATGCAAGCAAAA 426  
 Db 337 CCGTATGAAGAAACAAAGAACTTGTAAAGATATCTCAAAAGAGATGATGCAAGCAAAA 396  
 Oy 427 GATCTGACTGAGGCGAGTAGAGGTAAAGATGAAGATGCGATTTGAAATTTGAGCAAGAGC 486  
 Db 397 GATCTGACTGAGGCGAGTAGAGGTAAAGATGAAGATGCGATTTGAAATTTGAGCAAGAGC 456  
 Oy 487 ACTGTAAGGTCACAGGCGACACACAGAGATGTAAGCAGATTAAGCTTAATGATGAGG 546  
 Db 457 ACTGTAAGGTCACAGGCGACACACAGAGATGTAAGCAGATTAAGCTTAATGATGAGG 516  
 Oy 547 GTTCTGTTTGAAGAGCGCTTCTGAAGCAGAGAGATGAGAGCCCTTTGCAATAAC 606  
 Db 517 GTTCTGTTTGAAGAGCGCTTCTGAAGCAGAGAGATGAGAGCCCTTTGCAATAAT 576  
 Oy 607 GATTAAGTGTTCGCTGTTGCTTTCAGAGATTAAGACTCCCTTACTTTTGGGCTCCACGG 666  
 Db 577 GATTAAGTGTTCGCTGTTGCTTTCAGAGATTAAGACTCCCTTACTTTTGGGCTCCACGG 636  
 Oy 667 TTCTCTGCTCAATTTATGATGATCAAGTTCGCAAGAAATTAAGCTGATGATTTGATGAT 726  
 Db 637 TTCTCTGCTCAATTTATGATGATCAAGTTCGCAAGAAATTAAGCTGATGATTTGATGAT 696  
 Oy 727 GCCAAGGTTTGGAGAGAGCTTGAACATCAATGACACGATTCATGATTTGATGATCCTG 786  
 Db 697 GCCAAGGTTTGGAGAGAGCTTGAACATCAATGACACGATTCATGATTTGATGATCCTA 756  
 Oy 787 TGTGATGATGATTTGATGATGATCAAGTATCGGAGGCGCAAGAGCTGCAAACTT 846  
 Db 757 TGTGATGATGATTTGATGATGATCAAGTATCGGAGGCGCAAGAGCTGCAAACTT 816  
 Oy 847 ATTCTCAACATGAGTGCATGAAGAGATCTTGAGAAATCTTAATAAAGACAGATATCAA 906  
 Db 817 ATTCTCAACATGAGTGCATGAAGAGATCTTGAGAAATCTTAATAAAGACAGATATCAG 876  
 Oy 907 ATTCTGAGAGAGCTTACCAAGAGAGCTGAGCTTTCAAGGAGCTTAATGATCCT 966  
 Db 877 ATTCTGAGAGAGCTTACCAAGAGAGCTGAGCTTTCAAGGAGCTTAATGATCCT 936  
 Oy 967 TTGATATTTCTGAGCTAAATGCACTGACCTGATGAGAGAGGCTTCATTAAGTTTCTCG 1026  
 Db 937 TTGATATTTCTGAGCTAAATGCACTGACCTGATGAGAGAGGCTTCATTAAGTTTCTCG 906



Db 937 TTGGATTTCCTGACCTGAATGCAATGCTCCCGCATGAGAGGCTTGTAGATTCTG 996  
 QY 1027 GTAAAAAGATTAATGTTTCAATGAGATCGGGTACCAAGGCCATAGAGAGATCAATCT 1086  
 Db 997 GTGAAAGAAATGTTTCAACACAGATAGAGTACCAAGGCCATAGAGATTAATTT 1056  
 QY 1087 GCCAAGATTAATGCTGCGCAGAGAGACTCGAGTCTTTTTCACCAAGCCACCA 1146  
 Db 1057 GCCAAGAACCAATCTTCCCAAGAGGCTCGATCTCTTTCACCCAGTTGTAGCACA 1116  
 QY 1147 TCAGCACCCTAAACGAGAGAGACTTGCATTAACAGCAGCAGCTGCGAACAG 1206  
 Db 1117 TCAGTCCCTCTGAAAAAGAGACACTTCAAGAAAAACCAACCAAGCTGCTAACAG 1176  
 QY 1207 AAAACAAAG--GCTGGTGGAAAAAGAAATATCTTGTGATCTGTAGTACACTACGA 1263  
 Db 1177 AAAACAAAGGAGCGCGGCAAGAAAGAAATAGTGTGTGCTGCTATTCATCAACAAGA 1236  
 QY 1264 C 1264  
 Db 1237 C 1237

RESULT 6  
 XLU68141 1469 bp mRNA linear VRT 26-JUL-1999  
 LOCUS XLU68141 1469 bp mRNA, complete cds.  
 DEFINITION Xenopus laevis XFEN1b mRNA, complete cds.  
 ACCESSION U68141  
 VERSION U68141.1 GI:1549392  
 KEYWORDS  
 SOURCE African clawed frog.  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 1469)  
 Bibikova,M., Wu,B., Chi,E., Kim,K.H., Trautman,J.K. and Carroll,D.  
 Characterization of FEN-1 from Xenopus laevis. CDNA cloning and  
 role in DNA metabolism  
 J. Biol. Chem. 273 (51), 34222-34229 (1998)

REFERENCE  
 JOURNAL 99069415  
 MEDLINE 9852084  
 PUBMED 2 (bases 1 to 1469)  
 REFERENCE Bibikova,M., Chi,E., Wu,B., Kim,K.-H. and Carroll,D.  
 AUTHORS Direct Submission  
 JOURNAL Submitted (27-AUG-1996) Biochemistry, Univ. of Utah, 50 N. Medical  
 Drive, Salt Lake City, UT 84132, USA  
 FEATURES  
 source  
 1..1469  
 Location/Qualifiers  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 /clone\_lib="Hemmati-Brianlou et al. (1991) Development  
 111, 715-724"  
 /dev\_stage="tadpole"  
 /note="also present in oocytes, as determined by  
 sequencing of RT-PCR products"  
 120..1268  
 /function="5'-3' exonuclease and a structure-specific  
 endonuclease that removes 5' single-stranded 'flaps';  
 implicated in processing of Okazaki fragments during DNA  
 replication and in base excision repair; 5' nuclease"  
 /note="FEN-1 homolog; similar to mammalian Dnae IV or  
 MFL, to S. cerevisiae RAD27, RTH1 or YKL510, and to S.  
 pombe rad2; similar to Xenopus laevis XFEN1a encoded by  
 GenBank Accession Number U64563"  
 /codon\_start=1  
 /product="XFEN1b"  
 /protein\_id="AB08478.1"  
 /db\_xref="GI:1549393"  
 /translation="MGIHLAKLADVAPAIKREHDIKSYFGRKVAVDASMCITYOFLL  
 AYRQDGMIONEGETTHLMGFYFTIMLEIKPVVFPFGKPKMSGLARLSE  
 RRAAEKLEAAEEAGEVENTEFNRLVKVTKQNEECKLILSMGITVYVAFCEAE

ATCAALVRKAGVYAAATEMDALTEFGTPEVLLNHLTLASEAKRLPIQEFHLNRFQDIGI  
 NHEOFVDCILIGSDYCEITREIGPKRAIDLIFRMKEGEOFSDRIRNCAKLLANRQGS  
 KAROLPFLPEVYIDADITETKTEPDEGVIAFUMGCEKQFSDRIKXKYP1PENWHL  
 STQGRDDEFFKVTGSISTKRREVSSTKSKKSKTGGTGPACKEFRCK"  
 ORIGIN  
 BASE COUNT 461 a 288 c 377 g 343 t

Query Match 22.3%; Score 330; DB 5; Length 1469;  
 Best Local Similarity 58.3%; Pred. No. 3.7e-71;  
 Matches 617; Conservative 0; Mismatches 435; Indels 6; Gaps 2;

QY 90 AGACGAGATGGGATCAAGGGTTTGAGCAAACTGCTGGCGGACATGCGCCAGCGCAT 149  
 Db 113 AGCAATATGGAATTCACGCTTGCGCAAGCTTATCGCAGATGTGGCGCTGACACTAT 172  
 QY 150 GAAGGACAGAGATGTGAGAGCTACTTGGCGCAAAATGCGCGAGCGCCAGCATAG 209  
 Db 173 CAAAGACATGATATTAATAAGATTCTTGTCCAAAAGTGGCCGTTGATCTTATG 232  
 QY 210 CATCTACAGTTCCTGATAGTAGTTGAGAGAGAGAGAGCATGGAACCTCACAATGAGC 269  
 Db 233 CATTTACAGTTCCTTATGCTGTCACAGAGATGCA--ATATGCTGAGAAATGAGA 289  
 QY 270 TGTGAGTCACTAGTATTTCACAGCAATCTTCAACCGGACATTAAGATTACTGAGAG 329  
 Db 290 AGCGCAACCAAGCCATCTCAAGGTATGTTTACCGTACCATTCGATGCTGATGAGCA 349  
 QY 330 GGGAAATCAAGCCATTTATGTTTATGGCAAGCCCTCGATATGAGAAACAACAAT 389  
 Db 350 TGGCATCAACCCGTTTATGTTTATGTTGAGAGAGCCCTCAGATGAGTACAGTGAAGT 409  
 QY 390 TGTAAAGATATCTCAAAAGAGATGATGCAACCAAGATCTGATGAGCAGTAGAGT 449  
 Db 410 AGCCAAAGCGATGAGGTAGGAGGAGAGAGAAATGCTTGAAGCTCCGAGAGGCG 469  
 QY 450 AGGAGATTAAGATCGATGTAATAATGAGCAAGAGAGCTGTAAGGTGACCAAGGCA 509  
 Db 470 TGGAGAACTGAGAACATGAGAGTTCAATAAAGCTTGTCAAGTGCACCAAGAGCA 529  
 QY 510 CAAGAGATGTAAGAGACTATTAAAGACTTATGAGGGGTCTGTTGATAGGACACCTTC 569  
 Db 530 CAATGAAAGTGAAGAAATGTTAGCTGATGAGTATTCCTATGATGATGACACCTG 589  
 QY 570 TGAAGCAGACAGATGTGACGCCCTTTGATTAACGATTAAGGTGTTGCTGTGCTTC 629  
 Db 590 TGAAGCTGAGGCCATGTCGCTTGTAGTAAAGACAGGAAAGTGTATGATGACGAAC 649  
 QY 630 AGAAGATAGGACATGCTTGTGAGGAGTCCAGAGTCTGCTGATTAATGATGCC 689  
 Db 650 TGAAGACATGATGCTGACCTTGTGACCTGCGGTGCTGTAGGACACCTCAGAGAG 709  
 QY 690 AAGTCCAAAGAAATACCTGTGATGAGAAATTTGATGTCACAGGTTTGGACAGACTTGA 749  
 Db 710 TGAAGCTTAAAGAGTGCCTTATTCAGAGAGTTCATCAACCGCTTTTTCAGGACATCG 769  
 QY 750 ACTGACCATGACACCTGATTTGATTTGATGATCTGCTGATGATGATGATGATGATG 809  
 Db 770 CATCAATCAGCAAAATTTGATGATCTGTATCTGTGCTGGAAGTGTATTTGTGAAC 829  
 QY 810 CATCAAGATGATGAGGAGGAGCAACAGCTGTGAACCTTATGTCACATGAGTGCATAGA 869  
 Db 830 AATCCGTGCTGATGACCCMAAAGGCCATGACCTGATAGCCAGCATAGACTATGA 889  
 QY 870 AAGCATCTTGGAGAAATTTAATTAAGACATATCAAAATTCCTGAGAGCTGCTTACCA 929  
 Db 890 ACAGATCATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 949  
 QY 930 AAGAGCTGAGCTTGTTCACAGAGAGCTAATGTCA---CTTGGATATCTGAGCTAAA 986  
 Db 950 GGAGGCGACAGCGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1009  
 QY 987 ATGAGCTGACCTGATGAGAGAGGCTCTCATTAAGTTTCTGTTAAAGATTAATGCTTCAA 1046  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||







/db\_xref="taxon:8355"  
 /cell\_type="oocyte"  
 1. .1149  
 /gene="FEN1"  
 1. .1149  
 /gene="FEN1"  
 /function="Structure-specific nuclease"  
 /codon\_start=-1  
 /product="flap endonuclease 1"  
 /protein\_id="AAB8707.1"  
 /db\_xref="GI:2674207"  
 /translation="MGTHGLAKLADYAPPAIKENDIKSYFGKRVADISMCYQELI  
 AVQDSNTLQNEGETTSLMGAFYRTIRVVEHGIKPYVVDKRPQMSGELAKRE  
 RRAAEKLEAAEEAGEVENIEKTRKLVYVTKOHNEECKLLTLMGIPYVDAPEAE  
 ATCAALVKAGKYVAATEEDALTEPTVLRHLRHTASEAKRLPJOEFHLNRPVODIGI  
 THEQFDLCILIGSDYCEITKRGISPRALDIROHMTIDETINIDIKKYPVENMTH  
 KEAKHLEPEVVDITDITELKTEPEDEGIVANMCGEKPQSEPRIRNGAKKLKNNQ  
 STQGRLDPEFKVTSYSSSTRKEAESKSKKKAKTGCTPAGKFKRK"

BASE COUNT 352 a 247 c 306 g 244 t  
 ORIGIN

Query Match 20.8%; Score 307; DB 5; Length 1149;  
 Best Local Similarity 55.6%; Pred. No. 1.8e-65;  
 Matches 633; Conservative 0; Mismatches 500; Indels 6; Gaps 2;

QY 97 ATGGGATCAAGGCTTTGACGAACTCTGCGGACAAATGCGCCCAAGCGATGAGAG 156  
 Db 1 ATGGGAATTCAGCTTTGGCCAACTTATGCAAGATGTGCACCTGACGATCAAGAG 60  
 QY 157 CAGAGCTTCAGAGCTCTCTGCGGCAAAATGCGCCCTGACCGCAGCATGAGATCTAC 216  
 Db 61 CATGATTTAAAGTTACTTGTGCGAAGGTGGCGATCGATGCTTCATGTGATTTAT 120  
 QY 217 CAGTCTCTGATAGTACTTGGAGACAGCGATGGAATCTCACAATGAGAGTGTGAA 276  
 Db 121 CAGTCTCTTATGCGGAGACAGAGATGGCA--ACACGCTGCAAAAGAAAGATGAA 177  
 QY 277 GTGATAGTATTTGACAGAAATGTTCAACCGACATAAGATTACTGGAACGGGAATC 336  
 Db 178 ACCACAGCATCTTATGATGATGTTTACCGAACCATTTGCAATGGTAGAGCATGGCATC 237  
 QY 337 AACGATTTATTTTGTGAGCAAGCTCTGATATGAAGAACAAGATTTGCTAAA 396  
 Db 238 AAMCAGATATGTGTTGATGGGAACCCCCACAGATGAAGTCGGTGAATGACCAAA 297  
 QY 397 AGTACTCAAAAAGATGATGCAACCAAGATCTGACTAGGACATGAGATGAGAT 456  
 Db 298 CGCAGTAGCGTAGGCGCGAGCAAGAAAGTTGCTTGAAGCTGCGGAGAGCGTGGAGAA 357  
 QY 457 AAGATGCGATTTGAAAATGAGCAGAGAGACGTGTAAGTTCACAGGCAACACAGCAA 516  
 Db 358 GTGGAGACATTTAGAAATGCTTAAGAGCTTGAAGGTATACCAAGCAGACCAATGAG 417  
 QY 517 GATTGTAACGACTATTAAAGACTTATGGGGTCTCTGTTAGAGCAGCCTTGTGAAGA 576  
 Db 418 GAGTGAAGAAATTAATTAACCTAATGGGATTCCTTAATGATGACACCGTGTGAAGCT 477  
 QY 577 GAAGCAATGTGACACCTTTGCAATTAACGATTAAGGTGTTGCTGTTCAGAGAT 636  
 Db 478 GAGGCCACGTGTGCTCTTAGTAAAGAGAGGAAAGTGTATGTCAGACAGCAAGAAC 537  
 QY 637 AAGGATCTCTTCTTTGGGGCTCCAGGCTCTCTGTCATTAATGATGATCAAGTTCC 636  
 Db 538 ATGATGCCCCCTGACCTTTGGACCCCTGTGCTCTTAGGACCTCAAGAGAGAGCT 597  
 QY 697 AAGAAAATACCTGTGATGAAATTTGATGTTGCCAAGCTTTGGAGAGCTTGAACACAC 756  
 Db 598 AAAAAGCTGCTATTCAGAGATTCATTAACCGTGTATTAAGACATAGATATACACA 657  
 QY 757 ATGACACAGTCAATTTGATTTGATCTGCTGTGATGATGATGATGATGATGATGAT 816  
 Db 658 CATGAACAATTTGTGATCTGTGATCTTGTCTGGAAGTACTACTGTGAACCATCCGT 717

QY 817 GGTATCGGGGGCAACACAGCTCTGAACTTATGTCACATGGGTCCATGAAGACATC 876  
 Db 718 GGCAATGACCCCAAAAGAGCCATTGACTGATCCGTGACATTAAGACCATTAAGACATC 777  
 QY 877 TTGAGAAATCTTAATTAAGACAGATATCAAAATTCCTGAGACCTGACCTTACCAAGAACT 936  
 Db 778 ATTGATACATTTGACCTCAAGAAATTAACCCGCTCCGGAATACTGGCTCCACAGAGAGCA 837  
 QY 937 CGACCGTGTTCAGAGAGCTTAA--TGTCACTTTGGATATTTCTGAGCTTAAATGAGCT 993  
 Db 838 AAACATCTTTCTTGGAAACCGAAGTATGCAACATCCACATCAGGAGCTAAGTGGATC 897  
 QY 994 GCACCTGATGAGAGGCTCTATAGTTTCTGCTGTAATAAGATTAATGTTCAATGAAGAT 1053  
 Db 898 GAGCCCATGAAGAAGATTTGTTGCTTCAATGTGTGTTGGGAGAAAGACATTTAGTGAAGAT 957  
 QY 1054 CGGTCACAAAGGCCATGAGAGATCAAAATCTGCCAAGAAATTAATGCTCCCAAGAAAGA 1113  
 Db 958 CGCATAGCAACGGGGCCAAAGTTGGCCAAAGAACCGCAAGCAAGCAGACACAGGCCGG 1017  
 QY 1114 CTGAGTCTCTTTCAAGCCCACTGCCACCATCAGACCGCTTAAAGCGAAGAGACT 1173  
 Db 1018 CTGATGACTTTTCAAGATGATGATCGGATCGTTAGCTCACTAAAGAAAGAGCAGAA 1077  
 QY 1174 TCGGATAAACACAGCGCAGCTGCGAACAAAGAAACAAAGGCTGGTGGAAAGAGA 1232  
 Db 1078 TCCAAGAGATCCGCTAAAGAAAGCCAAACTGAGAGTACCCTCGAGGAAGTTCAA 1136

RESULT 10  
 BC010203 2163 bp mRNA linear ROD 22-OCT-2001  
 LOCUS  
 DEFINITION  
 Mus musculus, similar to flag structure-specific endonuclease,  
 clone MGC:7137 IMAGE:3158176, mRNA, complete cds.  
 ACCESSION  
 BC010203  
 VERSION  
 BC010203.1 GI:16307327  
 KEYWORDS  
 MGC.  
 SOURCE  
 house mouse.  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Strausberg, R.  
 Direct Submission  
 Submitted (05-JUL-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Jeffrey Green M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [villalobebcm.tmc.edu](mailto:villalobebcm.tmc.edu)  
 Villalon, D.K., Luna, R.A., Hale, S.M., Huilyk, S., Lu, X., Garcia,  
 A.M., Holloway, M., Yellford, B., Hodgson, A., Bouck, J., Yu, W.,  
 Muzny, D.M., Gibbs, R.A.

FEATURES  
 source  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAC Plates: 6 Row: d Column: 19.

Location/Qualifiers  
 1. .2163  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="MGC:7137 IMAGE:3158176"  
 /tissue\_type="Mammary tumor, C3(1)-Tag model, infiltrating  
 ductal carcinoma, 5 month old virgin mouse."

```

/cdclone.lib="NCL.CGAP_Mam6"
/lab_host="DHI0B"
/notes="Vector: PCMV-SPORT6"
206..1348
/codon_start=1
/product="Similar to flag structure-specific endonuclease"
/protein_id="AAH10203.1"
/ab_xref="GI:16307328"
/translation="MGIHQLAKLIVDAPSAIRENDIKSYGRKVAIDASITVFLI
AVRQGDVLONEGETTSHLMGMEYRTIMMENNGIKPYVFDGKPPQLSGELAKRSE
RAAEKEDLOQAEQMEVEDEKFTRLVKYKTHOHNDCKHLISLIPYDAPSEAE
ASCAALAKAGKVVAAATBEDMDLTFGSPVLMHLIOLKRSLEIYRTPADPKYVPEWHL
NOEFDYDLICILSDYCESIRGIEPRKAVLDLIDOKRSLEIYRTPADPKYVPEWHL
KEAOGLFLEPEVIDPESVELKSEPENEELVFMGKESQFSEERIRSGVKRLSNRQ
STQGRDDDFKVTGSLSSAKRKEPPEPKAKKATGCAAGAFRRKK"
BASE COUNT      540 a      528 c      619 g      476 t
ORIGIN
Query Match      20.6%; Score 303.8; DB 10; Length 2163;
Best Local Similarity 56.2%; Pred.No.1.3e-64;
Matches 641; Conservative 0; Mismatches 482; Indels 18; Gaps 3;
QY 97 ATGGGCAATCAAGGGTTTGGCAAACTGCTGGCGGACAAATGCGCCCAAGCGATGAAGAG 156
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 206 ATGGGAATTCAGCGCCCTGCCAAACTAATGCTGATGTGGCCCCCGCATGCAATCGGTAG 265
QY 157 CAGAAATTCAGAGACTACTTCGGCCGCAAAATGCGCGTGCAGCGCCGATGAGCATCTAC 216
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 266 AATGACATCAAGAGACTTGTGGTGCAGAAAGTGCGCATGATGCTTCACATGACATCTAC 325
QY 217 CAGTTCCTGATAGTATGTTGGAAGACAGCATGAAACTCTCAACAATGAAGCTGGTGAA 276
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 326 CAGTTCCTGATAGTATGTTGCTGTCAG ---GTTGGGAGTGTGCTGAGAAACGAGAGGGTGA 382
QY 277 GTCTACTAGTCAATTTGCAAGAAATGTTCAACCGCAATTAAGATTACTGCAAGCGGCAATC 336
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 383 ACCACCGACCACTGATGGGAGCATGTTCTACCGTACCATCCGATGATGGAATGTCATC 442
QY 337 AAGCCACTTATGTTTGTGATGAGCAAGCCCTCCGATATGAGAAACAAAGCACTTGTATA 396
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 443 AAGCCTGTGATGCTTGTGATGAGCAACCAACCACTGAGTCAAGCGCGCTGGCCAGAG 502
QY 397 AGATATCTCAAAAGAGATGATGCAACCAAGATCTGACTGAGCGAGTAGAGTAGAGAT 456
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 503 CGCAGTAGAGAGCGCGCGCGGAGGCTGAGAAAGCACTGAGAGGCTCAGAGAGGCTGGGATG 562
QY 457 AAAGATCGCATTTGAAAATTTGACAGAGAGACTGTAAAGTGCACAGGCAACCAACGAA 516
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 563 GAGGAGAGAGGTGGAGAGTTCCACCAAGAGGCTCGTGAAGGTCAACCAAGCAATGTAT 622
QY 517 GATTGTAAAGCACTATTAAAGCTTATGGGGGTCTCTGTTGTAGAGGCACTTCTGAAGCA 576
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 623 GAGTGCACCAACCTGCTGAGGCTCATATGGCATCTTACCTTATGACCCAGCAGAGGCA 662
QY 577 GAACAGAAATGTGCAGCGCTTTGCAATAAGGTGTTGCTGTTGCTTTCAGAGAT 656
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 683 GAGCCAGAGTGTGCTGCTGCGCAAGAGGTGCAAAAGTCTATGCTGCGCGCCAGAGAGAC 742
QY 637 AAGACATCCCTTACTTTTGGGGGCTCCAGGTTCTCTTGTGATTAATGAATCCAACTTCC 696
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 743 ATGATGATGCTTCACTTTTGGAGCCCGCTCTAATGCGACACTTAATCTGCACTGAGGCC 802
QY 697 AAGAAATATACCTGTGATGATTTGATGTTGCCAAGGTTTGGAGGAGCTGTAACCTCAC 756
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 803 AAGAAAGCTGCCATCCAAAGAGTTTCAATGACCGGCTCTGCAAGAGCTGGGGTCTGAAC 862
QY 757 ATGACAGAGTTCATGATTTGTGCATCTGTGTGATGTGATGATGATGATGATGATGATGAT 816
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 863 CAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 922
QY 817 GGTATCGGGGGGCAAGAGCTCTGAACCTTATTCGTAACATGGCTGCATGAGAAAGATC 876
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 923 GGCATTGTGGCCCAAGCGGGCTGTGATCTCATCGAAGCAATAGAGCATGAGAGATTC 982

```

```

QY 877 TTGGAGATCTTATTAAGACAGATATCAATTCCTGAGGACGTGCTTACCAAGAGCT 936
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 983 GTGAGCGCGCTGGAGACCCAGCATACCTCCCTTCCAGAGATGCTGCTCCACAGAGAGCC 1042
QY 937 CGAGCTGTGTCAAGAGAGCGCTTAATG---TCACCTTGGATATTCCTGAGCTTAATATGCACT 993
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1043 CAGAGAGCTCTTCGTGAGAGCCAGAGTACTGAGACCCAGAGCTGTGAGCTGAAAGTGAAGC 1102
QY 994 GCACCTGATGAGAGAGGCTCTCATAGTTTCTGCTGTTAAAGATAATGTTTCATGAAAT 1053
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1103 GAGCCAAATGAAAGAGAGTGTGCAAAATTTATGTGTGTGAAGAGCACTTTTCGAAGAG 1162
QY 1054 CGGGTGACAAAGGCCATGAGAGATGCAATCAATCTCCCAAGAAATAATGTCGACAGAGAGA 1113
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1163 CGAATTCGACAGTGGGATTAAGAGCGCTGATGATGAGAGCCGCCAGGACACCCAGAGAGCG 1222
QY 1114 CTCGAGTCCCTTTTTCAGCCCACTGCCACCATCAGACCGCTTAAGAGAGAGACT 1173
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 1223 CTCGATGATTTCTTCAAGAGTGAAGAGGCTCACTCTCTCAGC-----TAAACCGC 1270
QY 1174 TCGCATTAAGACAGAGGCGCTGCGACCAAGAAACAAAGGCTGTTGGAAGAAAGAA 1233
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1271 AAGGAGCCAGAAACCCAGAGGCGCTCTAAGAAAGCAAGCAAGACTGGGAGCGGGAAG 1330
QY 1234 T 1234
Db 1234 T 1234

```

```

RESULT 11
AC026761
LOCUS
DEFINITION
AC026761.10 GI:15217172
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 166088)
Gillis,G., Han,J., Montgomery,K.T., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
Mouse High Throughput Sequencing
Unpublished
2 (bases 1 to 166088)
Gillis,G., Han,J., Montgomery,K.T., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
Submitted (24-MAR-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
On Aug 21, 2001 this sequence version replaced gi:10937955.
-----Genome Center
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site:
http://twchanning.bwh.harvard.edu:9088/hpccg/jsp/hpccg/sequence/mous
e.html
Contact: gntm@capedoc.bwh.harvard.edu
-----Summary Statistics
Center project name: ACD
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye 100*
*Consensus quality: 156712 at least Q20
*Consensus quality: 152694 at least Q30
*Consensus quality: 147247 at least Q40
Estimated insert size: agarose-FP - N/A
*Estimated insert size: 165608 - sum-of-contrigs

```

COMMENT

# JOURNAL

```
Quality coverage: agarose-FP - N/A
Quality coverage: 4 x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 34556: contig of 34556 bp in length
* 3457 34576: gap of unknown length
* 62727 62727: contig of 28151 bp in length
* 62728 62747: gap of unknown length
* 62748 81216: contig of 18469 bp in length
* 81217 81236: gap of unknown length
* 81237 92870: contig of 11634 bp in length
* 92871 92890: gap of unknown length
* 92891 103627: contig of 10737 bp in length
* 103628 103647: gap of unknown length
* 103648 114697: contig of 11050 bp in length
* 114698 114717: gap of unknown length
* 114718 123385: contig of 8668 bp in length
* 123386 123405: gap of unknown length
* 123406 131815: contig of 8410 bp in length
* 131816 131835: gap of unknown length
* 131836 132037: contig of 202 bp in length
* 132038 132057: gap of unknown length
* 132058 138473: contig of 6416 bp in length
* 138474 138493: gap of unknown length
* 138494 144023: contig of 5530 bp in length
* 144024 144043: gap of unknown length
* 144044 144924: contig of 881 bp in length
* 144925 144944: gap of unknown length
* 144945 145191: contig of 247 bp in length
* 145192 145211: gap of unknown length
* 148414 148434: contig of 3203 bp in length
* 148435 151539: contig of 3105 bp in length
* 151540 151559: gap of unknown length
* 151560 154468: contig of 2909 bp in length
* 154469 154488: gap of unknown length
* 154489 156193: contig of 1705 bp in length
* 156194 156213: gap of unknown length
* 156214 157744: contig of 1531 bp in length
* 157745 157764: gap of unknown length
* 157765 159461: contig of 1697 bp in length
* 159462 159481: gap of unknown length
* 159482 160346: contig of 865 bp in length
* 160347 160366: gap of unknown length
* 160367 161824: contig of 1458 bp in length
* 161825 161844: gap of unknown length
* 161845 163663: contig of 1819 bp in length
* 163664 163683: gap of unknown length
* 163684 164031: contig of 348 bp in length
* 164032 164051: gap of unknown length
* 164052 164700: contig of 649 bp in length
* 164701 164720: gap of unknown length
* 164721 166088: contig of 1368 bp in length.
*
FEATURES
source
1. 166088
  /organism="Mus musculus"
  /strain="129/SVETACBr"
  /db_xref="taxon:10090"
  /chromosome="19"
  /clone="RP22-325J22"
  /sex="male"
misc_feature
1. 34556
  /note="assembly_name:Contig81"
misc_feature
34577..62727
  /note="assembly_name:Contig80"
misc_feature
62748..81216
```

```
misc_feature
81237..92870
  /note="assembly_name:Contig79"
misc_feature
92871..103627
  /note="assembly_name:Contig78"
misc_feature
103648..114697
  /note="assembly_name:Contig77"
misc_feature
114718..123385
  /note="assembly_name:Contig76"
misc_feature
123406..131815
  /note="assembly_name:Contig75"
misc_feature
131836..132037
  /note="assembly_name:Contig74"
misc_feature
132058..138473
  /note="assembly_name:Contig73"
misc_feature
138494..144023
  /note="assembly_name:Contig72"
misc_feature
144044..144924
  /note="assembly_name:Contig71"
misc_feature
144945..145191
  /note="assembly_name:Contig70"
misc_feature
145212..148414
  /note="assembly_name:Contig69"
misc_feature
148435..151539
  /note="assembly_name:Contig68"
misc_feature
151560..154468
  /note="assembly_name:Contig67"
misc_feature
154489..156193
  /note="assembly_name:Contig66"
misc_feature
156214..157744
  /note="assembly_name:Contig65"
misc_feature
157765..159461
  /note="assembly_name:Contig64"
misc_feature
159482..160346
  /note="assembly_name:Contig63"
misc_feature
160367..161824
  /note="assembly_name:Contig62"
misc_feature
161845..163663
  /note="assembly_name:Contig61"
misc_feature
163684..164031
  /note="assembly_name:Contig60"
misc_feature
164052..164700
  /note="assembly_name:Contig59"
misc_feature
164721..166088
  /note="assembly_name:Contig58"
misc_feature
166089..166088
  /note="assembly_name:Contig57"
BASE COUNT 44602 a 38872 c 39481 g 42573 t 560 others
ORIGIN
Query Match 20.6%; Score 303.8; DB 2; Length 166088;
Best Local Similarity 56.2%; Pred. No. 2.8e-64;
Matches 641; Conservative 0; Mismatches 482; Indels 18; Gaps 3;
11414 ATGGGCAATTCACGCGCTTGGCCAAACTAATTCATGTGAGCCGCCAGCCATCCGTCAG 11473
157 CAGAGTTCGAGAGCTACTTGGCCGCAAAATCCCGTCGAGCCAGCATGAGCATC 216
11474 AATGACATCAAGACTACTTGTGCGCAAACTGCGCATCGCTCTCATATGATCTAC 11533
217 CAGTTCCTGATAGTGTGGAAGGACGATGGAAGAACTCTCACAATGAAGCTGTGAA 276
11534 CAGTTCCTGATAGTGTGCGCAG---GTTGGGGATGTGTCGACAGAGAGAGGTGAG 11590
277 GTCACTAGTCATTGGCAAGGATGTTGACCGGACATAGATTAATGGAAGCGGAGATC 336
11591 ACCACCGCCACCTGATGGGCGATGTTTACCGTACCATCCGATGATGAGAAATGCGATC 11650
337 AAGCCATTATGTTTGTGATGGCAGCCTCTGATATGGAAGAAACAAGAACTTGTAAA 396
11651 AAGCCTGTACGCTTGTGATGGCAACCAACAGAGCTGAAGTGAAGGCGAGCTGCCAAG 11710
397 AGATACCAAAAAGATGATGACACCAAGATCTGACTGAGGCGATGAGAGTAGAGAT 456
```

```

Db 11711 CGCAGTATGAGGCGCCGCGGCTGAGAAAGCACTGACAGGCTCAGAGAGCTGGGAGT 11770
QY 457 AAGATGCGATTGAAAAATTGAGCAAGAGAGCTGTAAAGTGCACAAGCAACAACGAA 516
Db 11771 GAGAGAGAGGTGGAGAACTTACCAAGAGGCTGTGAGAGTACCAAGAACACATGTAT 11830
QY 517 GATTGTAACGACTATTAACTTATAGGGGCTTCTGTGTAGAGCACCCTTCTGAAGCA 576
Db 11831 GAGTGCACAAACCTGCTGAGACCTCATAGGGCATCCCTTACCTTATGATCACCACGAGCA 11890
QY 577 GAACAGATATGTGAGCGCTTTTGATTAACGATAGGTGTCTGCTGTGCTAGAGAT 636
Db 11891 GAGGCGAGCTGTGCTGCGCTGGCAAGGCTGCGCAAACTCATCTCTCGCGCCAGAGAGAC 11950
QY 637 AAGGACTCCCTTACTTTTGGGGCTCCACGGTCTCTGCTGCTTAATGATCCAAAGTCC 696
Db 11951 ATGGAAGCTCTCACTTTTGGACGCGCTGCTAATGCGACACTTAAGTCCAGTGAAGCC 12010
QY 697 AAGAAATACCTGTGATGATTTGATGATGTCAGAGTTTGGAGAGAGCTGAACACACC 756
Db 12011 AAGAAAGCTCCATCCAGAGTTCCATGAGCGCGCTCTGCAAGAGAGCTGGGTGAGC 12070
QY 757 ATGAGACAGTTCATGATTGATTCATCTGTCATCTGTGATGTGATGATGATGATGATCAAA 816
Db 12071 CAGAGAGAGTTTGTGATGTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12130
QY 817 GGTATCGGGGGGCAACAGCTCTGAACCTTATTCGTCACATGAGGCTCAGAGAGCATT 876
Db 12131 GGCATTTGGGGCCCAAGCGGCTGTGATCTCATCCAGAAACATTAAGACATGAGAGATC 12190
QY 877 TTGAGATATCTTATTAAGACAGATATCAATTCCTGAGAGAGAGGCTTACCAAGAGATC 936
Db 12191 GTGAGGCGGCTGGAGCCCAAGAGATACCCGCTTCCAGAGAGCTGCTCCAGAGAGAGCC 12250
QY 937 CGAGCTTTGTTCAAGAGACCTTAATG---TCATTTTGGATTTTCTGAGCTTAATAGGACT 993
Db 12251 CAGAGAGCTTCTCTGAGAGCAGAAATGCTGAGCCAGAGAGTCTGTGAGCTGAAGTGAAG 12310
QY 994 GCACCTATGAGAGAGGCTCATTAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1053
Db 12311 GAGCCAAATGAAGAAAGAGTGTCAATTTATGTTGTTAAAGCAAGTTTCTGTAAGAG 12370
QY 1054 CGGATGCAAAAGGCTATAGAGAGATCAATTCGCCAAGATTAATTCGTCGAAGAGAGA 1113
Db 12371 CGATTGCGAGTGGGGTCAAGCGGCTGATGAAGAGCGCGCAGGGCAGACACCGAGGACGC 12430
QY 1114 CTGAGTCTCTTTTCAAGCAACTGCCACATCAGCAGCGCTAAACGGAAGAGACT 1173
Db 12431 CTCGATATTTCTTCAAGGTGACAGGCTCCTCTCAGC-----TAAGCGC 12478
QY 1174 TCGATTAACCAACAGCAGCAGCTGCGAACAAGAAAACAAAGGCTGTGTGAAGAAAGAA 1253
Db 12479 AAGGAGCCAGAACCCAGAGGCGCTGCTTAAGAGAAAGCAAGAGAGGCGGAGAG 12538
QY 1234 T 1234
Db 12539 T 12539

```

```

RESULT 12
HSRAD2          1757 bp  mRNA  linear  PRI 07-DEC-1998
LOCUS           Homo sapiens mRNA for flap endonuclease-1.
DEFINITION      X76771.1 GI:1905802
VERSION         X76771.1
KEYWORDS        Ien1 gene; flap endonuclease-1.
SOURCE          human.
ORGANISM        Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1757)
AUTHORS        Watts,F.

```

```

TITLE           Direct Submission
JOURNAL         Submitted (16-DEC-1993) F. Watts, University of Sussex, School of
                Biological Sciences, Falmer, Brighton BN1 9QJ, UK
REFERENCE       2 (bases 1 to 1757)
AUTHORS         Murray,J.M., Tavassoli,M., al-Harithy,R., Sheldrick,K.S.,
                Lehmann,A.R., Carr,A.M. and Watts,F.Z.
TITLE           Structural and functional conservation of the human homolog of the
                Schizosaccharomyces pombe rad2 gene, which is required for
                chromosome segregation and recovery from DNA damage
                Mol. Cell. Biol. 14 (7), 4878-4888 (1994)
                94277093
FEATURES        location/Qualifiers
source          1..1757
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /gene="fen1"
                /gene="fen1"
                /gene="fen1"
                /codon_start=1
                /product="flap endonuclease-1"
                /protein_id="CA54166.1"
                /db_xref="GI:3980293"
                /translation="MGIOGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYOPLI
                AVROGQVLOMEGETSHLMGMYRTIRAMEGRIKPVYFDKRPOLKSGELAKRSE
                RRAERKLOAOAGAEOEVEKREKRLVTVTOHDECKHLISLGIPIYLDAPSAE
                ASCALVAKCKVYAATREMDCLTFGSPVIRHLTASEAKKLPIQEFHLSRIIOELG
                NOEOPVDICILISDICESIRGIGPKRAVDLIQKHSIEIVRLDPNKYPVENNLH
                KEAHQLEPEVLDPESEVLEKMSPEEELIKFMCGKQSEERISVGVARRLSKSHOG
                STQRLDPEFRVYGLSLSARKKEDEPFGSTTKRAKTKGAAGKFRGK"
BASE COUNT     418 a 443 c 542 g 354 t
ORIGIN
Query Match    20.3%. Score 299.6; DB: 9; Length 1757;
Best Local Similarity 54.8%; Pred. No. 1,4e-63;
Matches 659; Conservative 0; Mismatches 354; Indels 9; Gaps 3;
QY 97 ATGGCATCAAGGGTTTACGAAACTGCTGGCGGCAATGCGCCCAAGGCGCATGAAGAG 156
Db 368 ATGGCAATTCAGAGGCTGTGGCCAAACTAATTTGCTGATGTGGCCCAAGTCCCATCGGGAG 427
QY 157 CAGAAATTCAGAGCTACTTGGCGCCGCAAAATTCGCGCTGACGCCAGCATGACATCTAC 216
Db 428 AATGACATCAAGAGCTACTTGGCCGTAAGGTGGCCATGTGATGCTCATGATGAGCATTTAT 487
QY 217 CAGTCTCGATATGATGTGGAAGGACAGCATGGAATCTCACAAATGAAGCTGTGAA 276
Db 488 CAGTCTCGATATGCTGTGCGCAG---GGTGGGATGTGCTGCAAGATGAGAGGATGAG 544
QY 277 GTCACTAGTCATTTCAAGGATTTTCAACGAGCATGAATGAATTAAGTACTGAGAGCGGATC 336
Db 545 ACCACAGGACACTATAGGCGATGTTTACCGACACATTTCCATGATGAGAGAGGCGATC 604
QY 337 AAGCCAGTTTATGTTTTGATGCGCAAGCTCTCTGATATGAAGAAACAAGAACTTGCTAAA 396
Db 605 AAGCCCGTGTATGCTTTGATGCGCAAGCGCCACAGCTCAAGTCAAGGAGCTGCGCCAAA 664
QY 397 AGATATCTAAAAAGATGATGACCAACCAAGATCTGATGAGCGAGTGAAGTGAAGAT 456
Db 665 CCGAGTGAAGCGGCGGAGCTGAGGACAGAGAGCAAGCTGACAGCGAGCTCAGGCTGCGGGCC 724
QY 457 AAGATGCGATTGAAAAATTGAGCAAGAGAGCTGTAAAGTGCACAAGCAACAACGAA 516
Db 725 GAGCAGAGAGGTGAGAAATTTTCAAGCGGCTGTGGAAGGTCACCTAAGCAGACACAATGAT 784
QY 517 GATTGTAACGACTATTAACTTATAGGGGCTTCTGTGTAGAGCACCCTTCTGAAGCA 576
Db 785 GAGTGCACAAACCTGCTGAGAGCTCATAGGGCATCCCTTATCTTGAAGCACCAGTAGAGCA 844
QY 577 GAGCAGATATGAGCGCTTTGATTAACGATAGAGTGAAGTGTGCTGCTTCAAGAGAT 636
Db 845 GAGGCGAGCTGTGCTGCGCTGTGAGAGCTGCGCAAGATCTATGCTGCGGCTAACCGAGAGC 904

```

```

OY 637 AAGGACTCCCTTACTTTTGGGCTCCAGGCTTCCTGCTATTTAATGATCCAACTTCC 696
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 ATGGAGTCTGCTGCGAGCCCTGTGTAATGCGACACTGCTCCAGTGAAGCC 964
OY 697 AAGAAATACCTGATGGAATTTGATGTTGCCAAGTTTGGAGAGCTTGAACCTACCC 756
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 AAAAAAGTCCCATCAGGAATTTCCACCTGAGCCGGATTCTCAGAGAGCTGGGCTGAAC 1024
OY 757 ATGACAGTTCATTTGTCATTCCTGTTGGATGTCATTTGATGATAGCATMA 816
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1025 CAGGAACAGTTTGTGATCTGTGATCCGCTAGGACAGTACTGTGAGATATCCGG 1084
OY 817 GGTATCGGGGGGCAACAGCTCTGAATTTATTCCTCAACATGGTCCATAGAAAGCATC 876
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1085 GGTATTTGGCCCAAGCGGCTGTGACCTCATCCAGAAAGCACAGAGCATCAGAGAGATC 1144
OY 877 TTGGAATCTTAATAAAGACAGATATCAATTCCTGGAGAGCTTACCAAGATC 936
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1145 GTGGGCGAGCTGACCCCAACAGTACCTGTGCGAATAATTTGCTCCACAGAGGCT 1204
OY 937 CGAGCTTTGTTCAAGAGCCTAATG--TCACCTTGTGATATTCCTGAGCTAAATGACT 993
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1205 CACGAGCTCTTCTTGACACTGAGTGTGACCCAGAGTGTGTGGAGCTGAAGTGAGC 1264
OY 994 GCACCTGATGAGAGAGGCTCTATAGTTTCCGTGTAAGAATATGTTTCATGATAGAT 1053
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1265 GAGCCAAATGAAAGAGAGCTATCAATGTTGTTGTAAGAACAGTCTCTGAGAGG 1324
OY 1054 CGGGTGACAAAGGCGCATAGAGAAAGATCAAAATCTGCCAAGAAATATCGTCGCAAGAGAA 1113
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1325 CGAATCCGACAGTGGGGTCAAGAGGCTGAGTAAGAGCCGCCAAGGACAGCAGGAGGCGC 1384
OY 1114 CTCGAGTCTTTTCAAGCCACTGCGAC--CACATGACGACCCGCTAAAGCAAGAGAG 1170
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1385 CTGAGTGAATTTCTCAAGCTGACGGCTCACTCTCTCACTAAGCGAAGAGGACCAAA 1444
OY 1171 ACTTCGATTAACAAGAGGAGGCTGCGAACAAGAAACAAGGCTGTGGAAAGAG 1230
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1445 CCCAAGGATCTCACTAGAAAGAGCAAGAGCTGGGGCAGAGGGAAGTTTAAAGGGGA 1504
OY 1231 AATATATCTTGATGCTTGTATGATCACTAGCACTAGCAAGAGCGGTGCATGATCAG 1290
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1505 AATATATGTTTCTTCCCATATACCTCTTACCCAGAAATTTTGGCGTCTTACCC 1564
OY 1291 TT 1292
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1565 TT 1566
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
BC000323 2031 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, flap structure-specific endonuclease 1, clone
ACCESSION BC000323
VERSION MGC:8478 IMAGE:2821792, mRNA, complete cds.
KEYWORDS BC000323.1 GI:12653112
SOURCE MGC.
ORGANISM human.
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLES Mammalia; Eutheria; Primates; Catarrhini; Homiindae; Homo.
JOURNAL 1 (bases 1 to 2031)
Direct Submission
Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps@remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory

```

```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaitersburg, Maryland:
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Shevchenko, Y., Wetherby, R.W., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantirlop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 1 Row: 9 Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 1905802.

FEATURES
Location/Qualifiers
1..2031
/organism="Homo sapiens"
/db_xref="LocusID:2237"
/db_xref="taxon:9606"
/clone="MGC:8478 IMAGE:2821792"
/tissue_type="Lung, small cell carcinoma"
/clone_id="NIH-MGC-7"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
139..1281
/codon_start=1
/product="flap structure-specific endonuclease 1"
/protein_id="AAH00323.1"
/db_xref="GI:12653113"
/translation="MGIOGLAKLADVAAPSAIRENDIKSYRGRVAIDASNIYOFLL
AVRGGGVOLONEBETTSILMGMYFRTIRIMENQIKPYVDGKPPOLKSELKRSR
RRAEKOLOQAOAGAEOEVEKFTKRLVTKQHNDECKHLLSLMGIPYDASEAR
ASCAALVAKGVAAATVADMDCLTFEGSPVLRHLJASPAKLPLOEFLSRILDEIG
NOFOVDICLIVDCEISIRGTPRAVDLIDKHSLEIEVRLDPKRYVPENMLH
KEAHLDFLEPVDPEVELKMSPNBELIKFPCGGEKORSSEERIRSVKRLSKSRG
STGRIDDFRVTSLSSAKRKEPEPSTTKRKAKTGAAGFKRGR"

BASE COUNT 533 a 465 c 589 g 444 t
ORIGIN

Query Match 20.3% Score 299.6; DB 9; Length 2031;
Best Local Similarity 54.8%; Pred. No. 1.4e-63;
Matches 659; Conservative 0; Mismatches 534; Indels 9; Gaps 3;

OY 97 ATGGGCAATCAAGGCTTTGACGAACCTGCTGGCGGCAATGCGCCCAAGCGCATGAAGAG 156
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 ATGGGAATTTCAAGCGCTGCGCAACTAATTTGCTGATGTGGCCCCAGTCCGAGAG 198
OY 157 CAGAGTTTCGAGAGCTTCTTGGCCGCAAAATCGCGTGCAGCGCCAGCATGAGCATCTAC 216
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 AATGACATCAAGAGCTACTTTGGCGCTAAGCTGGCCATGATGCTTATGAGCATTTAT 258
OY 217 CAGTTCTGATAGTAGTAGTGAAGGACGAGCATNGAATCTCACAATAAGAGCTGTGTA 276
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 CAGTTCTGATAGTAGTAGTGAAGGACGAGCATNGAATCTCACAATAAGAGCTGTGTA 315
OY 277 GTCACTAGTCAATTTGCAAGAAATGTTCAACCGGACCAATTAAGATTACTGGAAGCGGAAATC 336
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 ACCACAGCCACACTGATGGGATGCTTACCGCAACATTCGATGATGAGAAAGCGGATC 375
OY 337 AAGCAGATTATGTTTGTGATGCGCAAGGCTCTGATATGAAGAAACAAGACTTGTCTAA 396
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 AAGCCCGTATGTTCTTGAAGGCAAGCGCCACAGCTCAAGCTCAGGAGGACCTGGCCAAA 435
OY 397 AGATACTCAAAAAGAGATGATCAACCAAGATCTGACTGAGAGAGTGAAGTGAAGAT 456
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 CGCAGTGACGGCGGGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 495
OY 457 AAGATGCGATTGAATAATTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516

```







\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```
1 20726: contig of 20726 bp in length
* 20727 20826: gap of 100 bp
* 20827 34666: contig of 13840 bp in length
* 34667 34766: gap of 100 bp
* 34767 46065: contig of 11299 bp in length
* 46066 46165: gap of 100 bp
* 46166 53516: contig of 7351 bp in length
* 53517 53616: gap of 100 bp
* 53617 61065: contig of 7449 bp in length
* 61066 61165: gap of 100 bp
* 61166 67998: contig of 6833 bp in length
* 67999 68098: gap of 100 bp
* 68099 74526: contig of 6428 bp in length
* 74527 74626: gap of 100 bp
* 74627 81299: contig of 6673 bp in length
* 81300 81399: gap of 100 bp
* 81400 87316: contig of 5917 bp in length
* 87317 87416: gap of 100 bp
* 87417 91541: contig of 4125 bp in length
* 91542 91641: gap of 100 bp
* 91642 95733: contig of 4090 bp in length
* 95732 95831: gap of 100 bp
* 95832 101128: contig of 5297 bp in length
* 101129 101228: gap of 100 bp
* 101229 104948: contig of 3720 bp in length
* 104949 105049: gap of 100 bp
* 105049 110306: contig of 5258 bp in length
* 110307 110406: gap of 100 bp
* 110407 112969: contig of 2563 bp in length
* 112970 113069: gap of 100 bp
* 113070 116484: contig of 3415 bp in length
* 116485 116584: gap of 100 bp
* 116585 119599: contig of 3015 bp in length
* 119600 119699: gap of 100 bp
* 119700 121881: contig of 2182 bp in length
* 121882 121981: gap of 100 bp
* 121982 122703: contig of 722 bp in length
* 122704 122803: gap of 100 bp
* 122804 125566: contig of 2763 bp in length
* 125567 125666: gap of 100 bp
* 125667 126845: contig of 1179 bp in length
* 126846 126945: gap of 100 bp
* 126946 129349: contig of 2404 bp in length
* 129350 129449: gap of 100 bp
* 129450 131268: contig of 1819 bp in length
* 131269 131368: gap of 100 bp
* 131369 132565: contig of 1197 bp in length
* 132566 132665: gap of 100 bp
* 132666 134031: contig of 1366 bp in length
* 134032 134131: gap of 100 bp
* 134132 136176: contig of 2045 bp in length
* 136177 136276: gap of 100 bp
* 136277 137418: contig of 1142 bp in length
* 137419 137518: gap of 100 bp
* 137519 138627: contig of 1109 bp in length.
```

## FEATURES

source

```
1.138627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="CMB9-22K5"
1..20726
/misc_feature
/note="assembly_fragment"
20827..34666
/misc_feature
/note="assembly_fragment"
34767..46065
/misc_feature
/note="assembly_fragment"
```

```
misc_feature 46166..53516
/note="assembly_fragment"
misc_feature 53617..61065
/note="assembly_fragment"
misc_feature 61166..67998
/note="assembly_fragment"
misc_feature 68099..74526
/note="assembly_fragment"
misc_feature 74627..81299
/note="assembly_fragment"
misc_feature 81400..87316
/note="assembly_fragment"
misc_feature 87417..91541
/note="assembly_fragment"
misc_feature 91642..95731
/note="assembly_fragment"
misc_feature 95832..101128
/note="assembly_fragment"
misc_feature 101229..104948
/note="assembly_fragment"
misc_feature 105049..110306
/note="assembly_fragment"
misc_feature 110407..112969
/note="assembly_fragment"
misc_feature 113070..116484
/note="assembly_fragment"
misc_feature 116585..119599
/note="assembly_fragment"
misc_feature 119700..121881
/note="assembly_fragment"
misc_feature 121982..122703
/note="assembly_fragment"
misc_feature 122804..125566
/note="assembly_fragment"
misc_feature 125667..126845
/note="assembly_fragment"
misc_feature 126946..129349
/note="assembly_fragment"
misc_feature 129450..131268
/note="assembly_fragment"
misc_feature 131369..132565
/note="assembly_fragment"
misc_feature 132666..134031
/note="assembly_fragment"
```

Query Match 20.3%, Score 299.6; DB 2; Length 138627;  
Best Local Similarity 54.8%; Pred. No. 3e-63;  
Matches 659; Conservative 0; Mismatches 534; Indels 9; Gaps 3;

```
QY 97 ATGGGCATCAAGGCTTTGACGAACCTGCTGGCGGACAAATGCGCCCAAGCGCATGAAGGAG 156
    ||||| || || || || || || || || || || || || || || || || || || || || ||
Db 17322 ATGGCAATTCGAAGGCTTGCCCAACTAATTCCTGATGTGGCCCCCAAGTCCGCGGAG 17381

QY 157 CAGAAGTTCGAGAGCTATTGGCGCGCAAAATCCGCGTCGACGCCAGCATAGCATCTAC 216
    || || || || || || || || || || || || || || || || || || || || || || ||
Db 17382 AATGNCATCAAGAGCTACTTTGGCCGTAAAGTGGCCATGATGCTCTATAGCAATTTAT 17441

QY 217 CAGTTCCTGATAGTACTTGAAGAGCATGAAGAACTCTCACAAATGAAGCTGTGTA 276
    ||||| ||||| || || || || || || || || || || || || || || || || || || ||
Db 17442 CAGTTCCTGATGCTGTTCGCCAG---GGTGGGATGTGCTGCGCAAAATGAGAGGAGGTAG 17498

QY 277 GTCACTACTCACTTTGCAAGCATGTTCAACCGGACCAATAGATTACTGGAAGCGGAATC 336
    || || || || || || || || || || || || || || || || || || || || || || ||
Db 17499 ACCACAGCCACCTGATGGGCGATGTTACCGCACCATTCGATGATGAGAGAGCGCATC 17558

QY 337 AAGCCAGTTTATGTTTGTGATGGCAAGCGCTCTGATATGAAGAACAAGACTTGTCTAA 396
    ||||| ||||| ||||| ||||| ||||| || || || || || || || || || || || || ||
Db 17559 AAGCCCGTATGTCTTTGATGAGCAAGCGCCACAGCTCAAGTCAAGGCGAGCTGGCCAAA 17618

QY 397 AGATAGTCAAAAAGATAGTGAACCAAGATCTGACTGAGCAGTAGAGTAGAGAT 456
    | ||||| || || || || || || || || || || || || || || || || || || ||
Db 17619 CCGAGTAGCGGCGGCTGAGAGCAGAGAGAGCTGCAGCAGGCTCAAGGCTGCTGGGGCC 17678

QY 457 AAAGATGCGATTGAATAATTGAGCAAGAGCACTTAAAGGTCACAGGCAACACACGAA 516
```

Db 17679 GAGCAGAGAGTGGAAAAATTCACCTAAGCGCGCTGGTGAAGTCTCTAGCAGCAATGAT 17738  
Qy 517 GATTGTAAAGACTATTAAAGCTTAAAGGGGGTCTGTTGTAGAGGACCTTCTGAAGCA 576  
Db 17739 GAGTGGAAACATCTCGCTGAGCCCTCATAGGGGATCCCTATCTTGATGACCCAGTAGGCA 17798  
Qy 577 GAAGCAGATATGTCAGCCCTTTCATTAAGCATTAAGTGTCTGCTTCTTCAAGAT 636  
Db 17799 GAGGCCAGCTGTCTGCTGCTGGTGAAGGCTGGCAAGATCATCTGCGCTACCGAGGAC 17858  
Qy 637 AAGGACCTCCCTTACTTTTGGGGCTCCAGGCTCTCTGCTGCTATTAATGATCCAAATGCT 696  
Db 17859 ATGACATGCTCTCAGCTTCCGAGCCCTGCTCTATGCGAAGCTGACTGCCAGTGAAGCC 17918  
Qy 697 AAGAAATACCTCTGATGATGATTTGATGTTGCCAAGGTTTGGAGAGCTTGAATCACC 756  
Db 17919 AAAAAGCTGCAATCCAGGAATTCACCTGAGCCGATCTCTGAGAGAGCTGGGCTGAAAC 17978  
Qy 757 ATGACAGATTCATGATTTGTGATGATGCTGATGATGATGATGATGATGATGATGATGAT 816  
Db 17979 CAGGAAAGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18038  
Qy 817 GGTATCGGGGGGCAAGACGCTGAAACTTATTCGTCAACATGAGTTCATAGAAAGCATC 876  
Db 18039 GGTATTTGGGGCCCAAGGGGCTGTGAGCTTATCCAGAGCAAGAGCATCGAGAGATC 18098  
Qy 877 TTGAGAAATCTTAAATAGACATATCAAAATTCCTGAGAGCTGGCTTACCAAGAGCT 936  
Db 18099 GTGGGGGAGTTGAGCCCAAGTACCTGCTGCTGCAAAATTTGCTCCCAAGGAGGCT 18158  
Qy 937 CGAGCTTTGTCAGAGAGCCCAATG---TCACCTTGTGATTTCTGAGCTTAAATGAGCT 993  
Db 18159 CACCAGCTCTTCTTGAACCTGAGTCTGAGCCAGAGCTCTGAGAGTGAAGTGAGC 18218  
Qy 994 GCACCTGATGAGAGGCTGCTCATAGTTTCTGCTAAAGATTAATGTTTCAATGAAGAT 1053  
Db 18219 GAGCAATATGAAGAAGAGCTGATCAAGTCTGCTGTAAGAGAGCTCTCTGAGAGAG 18278  
Qy 1054 CGGCTGCAAGAGCCATAGGAAGATCAATCTGCCAAGAAATTAATGCTGCCAAGAGAGA 1113  
Db 18279 CGAATCCGAGTGGGGTCAAGAGGCTGATGAGAGCCGCAAGGAGCAGCCAGGGCGC 18338  
Qy 1114 CTGCAATCTTTTCAAGCCCACTGCCAC---CACATCAGCAGCCGCTAAAGCGAAGAG 1170  
Db 18339 CTGATGATTTCTTCAAGAGTGAAGCGGCTCACTCTTCACTAGCAGGAGAGCCGAA 18398  
Qy 1171 ACTTCGATTAACACAGCAGGCTGCTGCAACAGAAAGCAAGAGCTGGTGAAGAGAG 1230  
Db 18399 CCAAGGATCCACTAAGAAAGAGCAAGAGCTGGGAGCAGAGGAGTTTAAAGGGGA 18458  
Qy 1231 AAATATCTTGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1290  
Db 18459 AAATTAATGTCTTCCCATATATACCTCTTCAACCCAGAAATATTTGCGCTTTGTACC 18518  
Qy 1291 TT 1292  
Db 18519 TT 18520

RESULT 15  
AP002380/c 169053 bp DNA linear HTG 31-MAY-2000  
LOCUS Homo sapiens chromosome 11 clone RP11-467L20 map 11q, WORKING DRAFT  
DEFINITION  
SEQUENCE 35 unordered pieces.  
ACCESSION AP002380.1 GI:8131644  
VERSION  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens DNA, clone:RP11-467L20.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 169053)

AUTHORS  
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y.  
TITLE  
JOURNAL  
Homo sapiens 169,053 genomic DNA of 11q  
Published Only in Database (2000) in press  
2 (bases 1 to 169053)  
REFERENCE  
AUTHORS  
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y.  
TITLE  
JOURNAL  
Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
Kiyasato Univ., 1-15-1 Kiyasato, Sagamihara, Kanagawa 228-8535,  
Japan (E-mail:hattori@gsc.riken.go.jp,  
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,  
Fax:81-42-778-9924)  
----- Genome Center  
COMMENT  
Center: RIKEN Genomic Sciences Center(GSC)  
Center code: RIKEN  
Web site: http://hgp.gsc.riken.go.jp/  
Contact: hattori@gsc.riken.go.jp  
----- Project Information  
Center project name: HumDrafl1  
Center clone name: RP11-467L20  
----- Summary Statistics  
Sequencing vector: PCR products; 100% of reads  
Chemistry: Dye-terminator ET-amersham; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 147827 bases at least Q40  
Consensus quality: 157936 bases at least Q30  
Consensus quality: 162410 bases at least Q20  
Insert size: 165653; sum-of-ctrls  
Quality coverage: 4.43x in Q20 bases; sum-of-ctrls  
-----  
NOTE: This is a 'working draft' sequence. It currently consists of  
35 contigs. The true order of the pieces is not known and their  
order in this sequence record is arbitrary. Gaps between the  
contigs are represented as runs 'N', but the exact sizes of the gaps  
are unknown. This record will be updated with the finished sequence  
as soon as it is available and the accession number will be  
preserved  
1 21322 contig of 21322 bp in length  
1 18823 contig of 18823 bp in length  
21423 40245 contig of 10147 bp in length  
40346 50593 contig of 7977 bp in length  
50593 58569 contig of 8716 bp in length  
58670 67486 contig of 7624 bp in length  
67486 75349 contig of 9105 bp in length  
75349 84453 contig of 9105 bp in length  
84453 89492 contig of 4939 bp in length  
89492 96445 contig of 6853 bp in length  
96445 102631 contig of 6086 bp in length  
102631 107300 contig of 4488 bp in length  
107300 111820 contig of 4521 bp in length  
111820 117068 contig of 5148 bp in length  
117068 120901 contig of 3733 bp in length  
120901 124072 contig of 3071 bp in length  
124072 127611 contig of 3479 bp in length  
127611 129401 contig of 1690 bp in length  
129401 129502 contig of 3640 bp in length  
129502 133141 contig of 3139 bp in length  
133141 136380 contig of 2945 bp in length  
136380 139425 contig of 3756 bp in length  
139425 143281 contig of 3068 bp in length  
143281 146448 contig of 2279 bp in length  
146448 148828 contig of 2117 bp in length  
148828 151045 contig of 2227 bp in length  
151045 153372 contig of 1524 bp in length  
153372 154996 contig of 1567 bp in length  
154996 156763 contig of 2269 bp in length  
156763 158132 contig of 2117 bp in length  
158132 161349 contig of 1401 bp in length  
161349 162850 contig of 1583 bp in length  
162850 164533 contig of 1196 bp in length  
164533 165829 contig of 1201 bp in length  
165829 167130 contig of 1060 bp in length  
167130 168290 contig of 1060 bp in length



Db 3399 GAGCAGAGAGCTGAAAAATTCACCTAAGCGGCTGTGAGGTCACTAAGCAGCACAATGAT 3340  
QY 517 GATTGTAAAGACATTAACCTTAAGGGGTTCTGTTAGAGGACCTTCTGAAGCA 576  
Db 3339 GAGTCAAAACATCTGCTGAGCCTCATGCGCATCCCTTATCTTGATGACCCAGTGAAGCA 3280  
QY 577 GAAGCAGAAATGTGAGCGCCCTTTGATTAAGCATAAAGTGTTCGCTTGCTTCAGAAAT 636  
Db 3279 GAGGCCAGCTGTGTGCTGCTGCTGGAAGGCTGGCAAGTCTATGCTGCGGTACCGAGGAC 3220  
QY 637 AAGGATCCCTTACTTTGGGGCTCCAGGTTCTTCCTCATTTAATGATCCAAGTCC 696  
Db 3219 ATGAGCTGCTCACTTCGGCAGCGCTGTGCTAATGACACCTGACTGCCAGTGAAGCC 3160  
QY 697 AAGAAATACCTGTGATGAAATTTGATGTCGAAGGTTTGGAGAGCTTGAACCTCAC 756  
Db 3159 AAAAAGCTGCCAATCCAGCAATTCACCTGAGCCGGATTTGCAAGAGCTGGGCTGAAAC 3100  
QY 757 ATGAGCAGATTCAATTGATTTGTGATCCTGTGATGTGATGATGATGATGATCAAA 816  
Db 3099 CAGGAACAGTTTGTGATGATGCTGCTGCTGAGGCACTGACTAGTGTGAGATATCCG 3040  
QY 817 GGTATCGGGGGGCAACAGCTCTGAACCTTATTCGTCAACATGGTCCATAGAAGCATC 876  
Db 3039 GGTATTTGGGCCCAACCGGCTGTGACCTCATCCAGAACCAAGCATCGAGGATC 2980  
QY 877 TTGAGATCTTATTAAGACAGATATCAATTCCTGAGGAGCTTCCCAAGAACT 936  
Db 2979 GTGCGGAGCTTGACCCCAACAGTACCTGTGTGCAAGAAATTTGGCTCCACAAGAGCT 2920  
QY 937 CGAGCTTGTTCAGAGAGCCTAATG--TCACTTGGATATTCCTGAGCTAAATGAGCT 993  
Db 2919 CAGCAGCTCTTCTTGAAACCTGAGTCTGAGCCACAGTCTGTGAGCTGAAGTGGAGC 2860  
QY 994 GCACCTGATGAGAGGCTCTCATTAAGTTCTGCTGAAGATTAATGTTCAATGAAGAT 1053  
Db 2859 GAGCCAATGAAGAGAGCTGATCAAGTTCATGTGTGTAAGAGCAGTTCTCTGAGGAG 2800  
QY 1054 CGGCTGACAAAGCCATAGAGAAATCAAAATCTGCCAAGAAATAAATGTCGAAGAGA 1113  
Db 2799 CGAATCCGCACTGGGCTCAGAGGCTGATAGAGCCGCCAAGGCAACACCAAGGCCG 2740  
QY 1114 CTCGAGTCTTTTTCAGAGCCAACTGCCAC---CACATCAGCACCGCTAAACGGAAGAG 1170  
Db 2739 CTGATGATTTCTTCAAGGTGACCGGCTCACTCTTTCAGCTAAGCCGCAAGGCCAGAA 2680  
QY 1171 ACTTCGATTAACACAGCAGGAGCTGCGAACAAGAAACAAAGGCTGCTGCAAGAGAG 1230  
Db 2679 CCGAAGGATCCACTAAGAAAGAGCAAGCTGGGCGACAGGAGAGTTTAAAGGGGA 2620  
QY 1231 AAATATCTTGATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1290  
Db 2619 AAATAAATGTGTTTCCCATTTATACCTCTTCAACCCCAAGATATTTGCCGTTGTACCC 2560  
QY 1291 TT 1292  
Db 2559 TT 2558

Search completed: November 5, 2002, 13:04:46  
Job time : 2474.66 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 08:10:48 : Search time 176.457 Seconds  
(without alignments)  
14380.822 Million cell updates/sec

Title: US-09-805-311-7

Perfect score: 1478  
Sequence: 1 cgaccacgcgtccgggaaa.....aaaaaaaaaaaaaaaaaaaa 1478

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: N.Geneseq\_032802:\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1478	100.0	1478	21	AAAZ7926
2	1421.8	96.2	1463	21	AAAZ7923
3	1401.2	94.8	1541	21	AAAZ7924
4	1334.2	90.3	1381	21	AAAZ7925
5	299.4	20.3	1144	20	AAAZ0217
6	299.4	20.3	1144	20	AAAZ02107
7	275	18.6	2033	20	AAAZ02111
8	274.4	18.6	1300	23	ABLI4287
9	270.6	18.3	1930	20	AAAZ02108

10	265.4	18.0	1149	20	AAAZ02109
11	251	17.0	3362	23	ABLI4286
12	173.2	11.7	5471	23	ABLI0094
13	162.8	11.0	1023	18	AAAZ76685
14	162.8	11.0	1023	19	AAAZ5840
15	162.8	11.0	1023	19	AAAZ3951
16	154.4	10.4	1054	19	AAAZ3984
17	152.8	10.3	1115	19	AAAZ3984
18	138.8	9.4	345980	22	AAAZ1224
19	137.2	9.3	1032	20	AAAZ1850
20	136.4	9.2	510	22	AAAZ37479
21	136.4	9.2	510	22	AAAZ18286
22	130.4	8.8	1115	19	AAAZ3978
23	129.6	8.8	514	19	AAAZ3990
24	128	8.7	1164	19	AAAZ3975
25	114.4	7.7	1729	23	AAAZ83732
26	111.8	7.6	889	19	AAAZ3974
27	100.8	6.8	1164	19	AAAZ3989
28	98.6	6.7	1011	19	AAAZ5913
29	98.6	6.7	1011	19	AAAZ54009
30	94.8	6.4	386	19	AAAZ3979
31	92.4	6.3	981	18	AAAZ76682
32	92.4	6.3	981	19	AAAZ3948
33	90	6.1	546	22	ABAZ3065
34	90	6.1	546	22	AAAZ11490
35	90	6.1	546	22	AAAZ37269
36	90	6.1	546	22	AAAZ18098
37	89.2	6.0	981	19	AAAZ5837
38	85	5.8	245	22	AAAZ50354
39	85	5.8	245	22	AAAZ12416
40	79.4	5.4	777	19	AAAZ54010
41	77.8	5.3	840	19	AAAZ5977
42	72.2	4.9	987	19	AAAZ54011
43	69.4	4.7	296	19	AAAZ3976
44	64.4	4.4	514	19	AAAZ3985
45	64	4.3	4551	21	AAAZ70129

## ALIGNMENTS

AAAZ7926	standard: cDNA; 1478 BP.
AAAZ7926	
AAAZ7926	
12-SEP-2000	(first entry)
Maize Rad2/FEN-1 CDNA.	
Maize: Rad2/FEN-1; transgenic plant; male sterile plant;	
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.	
OS Zea mays.	
XX	
FH	Location/Qualifiers
FT	CDS 97..1236
XX	/ftag= a
XX	
PN	WO200036109-A1.
XX	
PD	22-JUN-2000.
XX	
PF	16-NOV-1999; 99WO-US27147.
XX	
PR	15-DEC-1998; 98US-01123332.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Mahajan PB;
XX	
DR	Drosophila melanog
	WPI; 2000-452026/39.

Yeast FEN-1 CDNA.  
Drosophila melanog  
Drosophila melanog  
Drosophila melanog  
Pyrococcus furiosus  
Pyrococcus furiosus  
Nucleotide sequenc  
P. furiosus N-term  
P. furiosus N-term  
Pyrococcus abyssal  
Pyrococcus heat re  
Human bone marrow  
Probe #8219 for ge  
P. furiosus N-term  
Nucleotide sequenc  
Nucleotide sequenc  
DNA encoding novel  
Nucleotide sequenc  
M. jannaschii N-te  
Archaeoglobus fulg  
Nucleotide sequenc  
Nucleotide sequenc  
Methanococcus jann  
Nucleotide sequenc  
Human foetal liver  
Human brain expres  
Human bone marrow  
Probe #8031 for ge  
Methanococcus jann  
Human bone marrow  
Probe #17349 for g  
Nucleotide sequenc  
Nucleotide sequenc  
Nucleotide sequenc  
Nucleotide sequenc  
Nucleotide sequenc  
Plasmodium falcipar

DR P-PSDB: AAY95310.

XX Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA  
PT recombination and repair in transgenic plants, e.g. for gene targeting  
PT and the production of male sterile plants -

XX Example 1; Page 77-79; 85pp; English.

XX The present sequence is that of maize cDNA coding for Rad2/FEN-1  
CC (see AAY95310). The cDNA was isolated from a library prepared  
CC from B73 line seed vitreous endosperm RNA. Rad2/FEN-1 is a  
CC structure specific endonuclease which under certain conditions also  
CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to  
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.  
CC The protein is involved in the regulation of DNA repair and  
CC recombination in plant systems and therefore may be used for  
CC improving gene targeting during further recombinant DNA protocols  
CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential  
CC in DNA replication and nucleotide excision and repair reactions.  
CC The exolytic activity is involved in double strand break repair and  
CC end joining. The protein is also useful in strand exchange  
CC reactions during homologous recombination. These functions may be  
CC useful in gene targeting and in the production of male sterile  
CC plants. The efficacy of gene targeting can be improved by the  
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
CC be produced by the down regulation of Rad2/FEN-1 expression.

XX Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other;

Query Match 100.0%; Score 1478; DB 21; Length 1478;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACCCAGCGCTCGGGAATAGTCTCGGCTGCTTCTGCGCAGCTCGGCTCAGC 60  
DB 1 CGACCCAGCGCTCGGGAATAGTCTCGGCTGCTTCTGCGCAGCTCGGCTCAGC 60  
QY 61 CGCGCCCGCCAGCCGCGCAGCCGCGCAGAGAGATGGGATCAAGGTTTGAGGAA 120  
DB 61 CGCGCCCGCCAGCCGCGCAGCCGCGCAGAGAGATGGGATCAAGGTTTGAGGAA 120  
QY 121 CTGCTGGGGGCAATGGGCGCAAGCGATGAGAGAGATGTCAGAGCTTCTCGGC 180  
DB 121 CTGCTGGGGGCAATGGGCGCAAGCGATGAGAGAGATGTCAGAGCTTCTCGGC 180  
QY 121 CTGCTGGGGGCAATGGGCGCAAGCGATGAGAGAGATGTCAGAGCTTCTCGGC 180  
DB 121 CTGCTGGGGGCAATGGGCGCAAGCGATGAGAGAGATGTCAGAGCTTCTCGGC 180  
QY 181 CGCAAAATCGCGCTGAGCGCGCAGCATGAGCATCTACAGTTCTCTGATAGTTGGAAG 240  
DB 181 CGCAAAATCGCGCTGAGCGCGCAGCATGAGCATCTACAGTTCTCTGATAGTTGGAAG 240  
QY 241 ACAGGCAATGGAATCTTCACAAATGAAAGCTGTAAGTCTAGTATTTGCAAGAAATG 300  
DB 241 ACAGGCAATGGAATCTTCACAAATGAAAGCTGTAAGTCTAGTATTTGCAAGAAATG 300  
QY 301 TTCAACCGGCAATATGATTTACTGGAAGCGGAAATCAAGCAGTTTATGTTTGAATGCG 360  
DB 301 TTCAACCGGCAATATGATTTACTGGAAGCGGAAATCAAGCAGTTTATGTTTGAATGCG 360  
QY 361 AAGCTCTCTGATATGAGAAACAGAACTTCTTAAAGATCTCAAAAAGAGATGATGA 420  
DB 361 AAGCTCTCTGATATGAGAAACAGAACTTCTTAAAGATCTCAAAAAGAGATGATGA 420  
QY 421 ACCAAAGATCTGATGAGAGCGAGTAGAGTAGAGATTAAGATCGATTGAAAATGAGC 480  
DB 421 ACCAAAGATCTGATGAGAGCGAGTAGAGTAGAGATTAAGATCGATTGAAAATGAGC 480  
QY 481 AAGAGGAGCTGTAAGGCTCACAAGCAACACACAGAGATTTGTAACGACTTATTAAGCTT 540  
DB 481 AAGAGGAGCTGTAAGGCTCACAAGCAACACACAGAGATTTGTAACGACTTATTAAGCTT 540  
QY 541 ATGGGGGTTCTGTTGATAGAGCAGCTTCTGAAGCAGACAGAAATGTGAGCCCTTTGC 600  
DB 541 ATGGGGGTTCTGTTGATAGAGCAGCTTCTGAAGCAGACAGAAATGTGAGCCCTTTGC 600

QY 601 ATAAAGATTAAGGTTTCGCTGTTGCTCAGAGATTAAGAGCTCCCTTTTGGGCT 660  
DB 601 ATAAAGATTAAGGTTTCGCTGTTGCTCAGAGATTAAGAGCTCCCTTTTGGGCT 660  
QY 661 CCACGGTTCCTGCTCAATTAATGATCCAACTTCCAAAGAAATTCCTGTGATGAATTT 720  
DB 661 CCACGGTTCCTGCTCAATTAATGATCCAACTTCCAAAGAAATTCCTGTGATGAATTT 720  
QY 721 GATGTTCCCAAGGTTTGGAGAGCTTGAAGTCACTGACATGACAGTTCATGATTTGTC 780  
DB 721 GATGTTCCCAAGGTTTGGAGAGCTTGAAGTCACTGACATGACAGTTCATGATTTGTC 780  
QY 781 ATCTGTGTGATGTGATCTTATTTGATAGATCAAGATTCGAGGAGCAACAGCTCTG 840  
DB 781 ATCTGTGTGATGTGATCTTATTTGATAGATCAAGATTCGAGGAGCAACAGCTCTG 840  
QY 841 AAATTTTGTGCAACATGAGTGGCTCATGAAAGATCTTGAAGATCTTAAAGACAGA 900  
DB 841 AAATTTTGTGCAACATGAGTGGCTCATGAAAGATCTTGAAGATCTTAAAGACAGA 900  
QY 901 TATCAAAATTCCTGAGAGCTGCTTACCAAGAGCTGACGCTTGTCAAGAGCTTAAT 960  
DB 901 TATCAAAATTCCTGAGAGCTGCTTACCAAGAGCTGACGCTTGTCAAGAGCTTAAT 960  
QY 961 GTCACTTTGGATTTCTCTGAGTAAATGAGTCACTGACCTGATGAGAGAGGCTCATAGT 1020  
DB 961 GTCACTTTGGATTTCTCTGAGTAAATGAGTCACTGACCTGATGAGAGAGGCTCATAGT 1020  
QY 1021 TTCTGTGTAAGAAATGATTTGCTCAATGAAGATCGGGTGAACAAAGCCATAGAGAGATC 1080  
DB 1021 TTCTGTGTAAGAAATGATTTGCTCAATGAAGATCGGGTGAACAAAGCCATAGAGAGATC 1080  
QY 1081 AAATCTGCCAAGATTAATGCTGTCGCAAGAGAGCTGATGCTTTTCAAGCACTATGC 1140  
DB 1081 AAATCTGCCAAGATTAATGCTGTCGCAAGAGAGCTGATGCTTTTCAAGCACTATGC 1140  
QY 1141 ACCACATCAGACCGCTTAAACGGAAGAGACTTCGATTAACCAAGCAAGCAGCTGCG 1200  
DB 1141 ACCACATCAGACCGCTTAAACGGAAGAGACTTCGATTAACCAAGCAAGCAGCTGCG 1200  
QY 1201 AACCAAGAAACAAAGCTGCTGGAAGAAAGAAATTAATCTGATGCTTATACACTA 1260  
DB 1201 AACCAAGAAACAAAGCTGCTGGAAGAAAGAAATTAATCTGATGCTTATACACTA 1260  
QY 1261 CGACTACGAAAGCAGCGGTGATGATCACTTGCCTTAATTAATTAATCCCTGTTTGA 1320  
DB 1261 CGACTACGAAAGCAGCGGTGATGATCACTTGCCTTAATTAATTAATCCCTGTTTGA 1320  
QY 1321 ACTCAGACCTTTGGTGAAGTTTGGCCATGTTCAAGCTGGGGTAACTAGTTGTTTG 1380  
DB 1321 ACTCAGACCTTTGGTGAAGTTTGGCCATGTTCAAGCTGGGGTAACTAGTTGTTTG 1380  
QY 1381 AAGAGATTTGTTGTTCAAGATTAACAAATCTATGCTGTTTACTTGTGCTTTGAA 1440  
DB 1381 AAGAGATTTGTTGTTCAAGATTAACAAATCTATGCTGTTTACTTGTGCTTTGAA 1440  
QY 1441 GTATGATGCGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1478  
DB 1441 GTATGATGCGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1478

RESULT 2  
ID AAA27923 standard; cDNA; 1463 BP.  
XX AAA27923:  
AC 12-SEP-2000 (first entry)  
XX  
XX Maize Rad2/FEN-1 cDNA.  
XX  
XX Maize: Rad2/FEN-1; transgenic plant; male sterile plant;  
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.

XX OS Zea mays. Location/Qualifiers  
XX Key 85..1224  
XX FT CDS /\*lag- a  
XX FT  
XX PN MO200036109-A1.  
XX PD  
XX PD 22-JUN-2000.  
XX PF 16-NOV-1999; 99MO-US27147.  
XX PR 15-DEC-1998; 980S-0112332.  
XX PA (PION-) PIONEER HI-BRED INT INC.  
XX PI Mahajan PB;  
XX DR WPI; 2000-452026/39.  
XX DR P-PSDB; AAY95307.  
XX PT Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA  
XX PT recombination and repair in transgenic plants, e.g. for gene targeting  
XX PT and the production of male sterile plants -  
XX PS Claim 1; Page 69-71; 85pp; English.  
XX CC The present sequence is that of maize cDNA coding for RAD2/FEN-1  
XX CC (see AAY95307). The corresponding RNA was isolated from immature  
XX CC ear tissue from 2 ears of a B73 maize line. Rad2/FEN-1 is a  
XX CC structure specific endonuclease which under certain conditions also  
XX CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to  
XX CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.  
XX CC The protein is involved in the regulation of DNA repair and  
XX CC recombination in plant systems and therefore may be used for  
XX CC improving gene targeting during further recombinant DNA protocols  
XX CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential  
XX CC in DNA replication and nucleotide excision and repair reactions.  
XX CC The exolytic activity is involved in double strand break repair and  
XX CC end joining. The protein is also useful in strand exchange  
XX CC reactions during homologous recombination. These functions may be  
XX CC useful in gene targeting and in the production of male sterile  
XX CC plants. The efficacy of gene targeting can be improved by the  
XX CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
XX CC be produced by the down regulation of Rad2/FEN-1 expression.  
XX SQ Sequence 1463 BP; 466 A; 292 C; 361 G; 344 T; 0 other;  
Query Match 96.2%; Score 1421.8; DB 21; Length 1463;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 319 TTACTGAGACCGGAGATCAACAGCAGTTTATGTTTGTATGAGCAAGCCCTGATATGAG 378  
|||||  
Db 307 TTACTGAGAGGGGAAATCAAGCCAGTTTATGTTTGTATGAGCAAGCCCTGATATGAG 366  
QY 379 AAACAGAACTTGTCTAAAGATCTCAAAAAGAGATGATGACCAAAAGATGACTGAG 438  
|||||  
Db 367 AAACAGAGCTTGTCTAAAGATCTCAAAAAGAGATGATGACCAAAAGATGACTGAG 426  
QY 439 GCATAGAGTATGAGAGATTAAGATGCGATTGAAAATTTGAGCAAGAGAGCTGTAAAGTC 498  
|||||  
Db 427 GCATAGAGTATGAGAGATTAAGATGCGATTGAAAATTTGAGCAAGAGAGCTGTAAAGTC 486  
QY 499 ACAAGCAACACAGCAAGAGATTTGAAACGACTATTAAGACTATGAGGGGCTCCGTGTA 558  
|||||  
Db 487 ACAAGCAACACAGCAAGAGATTTGAAACGACTATTAAGACTATGAGGGGCTCCGTGTA 546  
QY 559 GAGGACCTTCTGAGAGCAAGAGAGATGTCAGACCCCTTGCATAAAGATTAAGGTTC 618  
|||||  
Db 547 GAGGACCTTCTGAGAGCAAGAGAGATGTCAGACCCCTTGCATAAAGATTAAGGTTC 606  
QY 619 GCTGTGCTTCAGAGAGATTAAGAGACTCCCTTACTTTGGGGCTCCAGGGTCTTCAT 678  
|||||  
Db 607 GCTGTGCTTCAGAGAGATTAAGAGACTCCCTTACTTTGGGGCTCCAGGGTCTTCAT 666  
QY 679 TTAATGATCCAACTTCCAGAAATACCTGTGATGATTTGATTTGATGATTTG 738  
|||||  
Db 667 TTAATGATCCAACTTCCAGAAATACCTGTGATGATTTGATTTGATGATTTG 726  
QY 739 GAGGAGCTTGAAGTACCAATGAGAGAGAGTTCATTTGATTCCTGTGATGAG 798  
|||||  
Db 727 GAGGAGCTTGAAGTACCAATGAGAGAGAGTTCATTTGATTCCTGTGATGAG 786  
QY 799 TATTGTATGATGATCAAGAGATTCGAGGAGGAGCAAGAGAGTTCATTTGATTC 858  
|||||  
Db 787 TATTGTATGATGATCAAGAGATTCGAGGAGGAGCAAGAGAGTTCATTTGATTC 846  
QY 859 GGGTCCATGAGAGAGATTCGAGAGATTCATTAAGAGAGATTCATTTGATTC 918  
|||||  
Db 847 GGGTCCATGAGAGAGATTCGAGAGATTCATTAAGAGAGATTCATTTGATTC 906  
QY 919 TGGCTTACCAAGAGAGTTCGAGAGAGTTCATTAAGAGAGATTCATTTGATTC 978  
|||||  
Db 907 TGGCTTACCAAGAGAGTTCGAGAGAGTTCATTAAGAGAGATTCATTTGATTC 966  
QY 979 GAGTAAATGAGAGAGTTCGAGAGAGTTCATTAAGAGAGATTCATTTGATTC 1038  
|||||  
Db 967 GAGTAAATGAGAGAGTTCGAGAGAGTTCATTAAGAGAGATTCATTTGATTC 1026  
QY 1039 GGTTCATGAGAGAGTTCGAGAGAGTTCATTAAGAGAGATTCATTTGATTC 1098  
|||||  
Db 1027 GGTTCATGAGAGAGTTCGAGAGAGTTCATTAAGAGAGATTCATTTGATTC 1086  
QY 1099 TCGTTCGAGAGAGAGTTCGAGAGAGTTCATTAAGAGAGATTCATTTGATTC 1158  
|||||  
Db 1087 TCGTTCGAGAGAGAGTTCGAGAGAGTTCATTAAGAGAGATTCATTTGATTC 1146  
QY 1159 AAAGGAGAGAGAGTTCGAGAGAGTTCATTAAGAGAGAGTTCATTTGATTC 1218  
|||||  
Db 1147 AAAGGAGAGAGAGTTCGAGAGAGTTCATTAAGAGAGAGTTCATTTGATTC 1206  
QY 1219 GGTTCGAGAGAGAGTTCGAGAGAGTTCATTAAGAGAGAGTTCATTTGATTC 1278  
|||||  
Db 1207 GGTTCGAGAGAGAGTTCGAGAGAGTTCATTAAGAGAGAGTTCATTTGATTC 1266  
QY 1279 TGGCATGATGATTCGAGAGAGTTCATTAAGAGAGTTCATTTGATTC 1338  
|||||  
Db 1267 TGGCATGATGATTCGAGAGAGTTCATTAAGAGAGTTCATTTGATTC 1326  
QY 1339 AGTTTGGCCATGTTTCAAGCTGGGGTATGATTTGATTTGATTTGATTTGATTTG 1398  
|||||  
Db 1327 AGTTTGGCCATGTTTCAAGCTGGGGTATGATTTGATTTGATTTGATTTGATTTG 1386  
QY 1399 GTACCAAACTTATCGCTGTTTTTACTTCTGCTTTGAGAGATGATGATGATGATGAT 1458  
|||||

Query Match	94.8%	Score 1401.2	DB 21	Length 1541
Best Local Similarity	99.1%	Ped. No. 0	Mismatches 13	Indels 0
Matches 1409	Conservative 0			Gaps 0

QY	33	CGCGTTTC	TTGGCCACTCGCGCTCAGGCCGCCGCCGCCACCCGCCACAGCCGCCGACGA	92
Db	15	GGGGTTCT	TGGCCCACTCGCGCTCAGGCCGCCGCCGCCACCCGCCACAGCCGCCGACGA	74
OY	93	CGAATGGG	CATCAAGGGTTTGACAAACCTGCTGGGGGCAATATGGCCCAAGGGCATGAA	152
Db	75	CGAATGGG	CATCAAGGGTTTGACAAACCTGCTGGGGGCAATATGGCCCAAGGGCATGAA	134
OY	153	GGAGCAGA	AGTTCGAGAGCTACTTGGGGCGCAAAATCGCGCTGCAGCGCAGATGAGCAT	212
Db	135	GGAGCAGA	AGTTCGAGAGCTACTTGGGGCGCAAAATCGCGCTGCAGCGCAGATGAGCAT	194
OY	213	CTACACCT	CTCTGATATTTGATTTGGAAGGACAGCATGGAACCTTCACAAATGGAAGCTGG	272
Db	195	ATACCACT	CTCTGATATTTGATTTGGAAGGACAGCATGGAACCTTCACAAATGGAAGCTGG	254
OY	273	TGAAGTC	ACTAGTCATTTGCAAGGAATGTTCAACCGGACATATAGATTACTGGAGCGGG	332
Db	255	TGAAGTC	ACTAGTCATTTGCAAGGAATGTTCAACCGGACATATAGATTACTGGAGCGGG	314
OY	333	AATCAACC	CCAGTTATGTTTTTATGGCAACCTCTCGATATGAGAAACAAAGACTGCG	392
Db	315	AATCAACC	CCAGTTATGTTTTTATGGCAACCTCTCGATATGAGAAACAAAGACTGCG	374
OY	393	TAAAGAT	ACACAAAGAGATGATCAACCAAGATCTGCATGAGGGCAGTAGAGTAGG	452
Db	375	TAAAGAT	ACACAAAGAGATGATCAACCAAGATCTGCATGAGGGCAGTAGAGTAGG	434
OY	453	AGATPAA	AGATGCGATTGAAAAATTGACCAAGAGGACTGTAAAGGTCACAAAGCAACACA	512
Db	435	AGATPAA	AGATGCGATTGAAAAATTGACCAAGAGGACTGTAAAGGTCACAAAGCAACACA	494
OY	513	CGAAGAT	TGTAAAGACTATTAAAGCTATAGGGGGTTCCTGTTGTAAGGGCACCTTCGA	572
Db	495	CGAAGAT	TGTAAAGACTATTAAAGCTATAGGGGGTTCCTGTTGTAAGGGCACCTTCGA	554
OY	573	AGCAGA	AGCAGATGTGCAGCCCTTTCATATAACGATTAAGGTGTTCGTTCCTTCAGA	632
Db	555	AGCAGA	AGCAGATGTGCAGCCCTTTCATATAACGATTAAGGTGTTCGTTCCTTCAGA	614
OY	633	AGATPAA	GACACCTTCCTTTGTTGGGGCTCCAGGGTTCCTCGCATTTAATGATCCAG	692
Db	615	AGATPAA	GACACCTTCCTTTGTTGGGGCTCCAGGGTTCCTCGCATTTAATGATCCAG	674
OY	693	TTCCAA	AGAAAAATACCTGTGATGGAATTTGTTGCCAAGGTTTGGAGGACCTTGAAC	752
Db	675	TTCCAA	AGAAAAATACCTGTGATGGAATTTGTTGCCAAGGTTTGGAGGACCTTGAAC	734
OY	733	CACCAT	GGACCATTTATTTGTCATACCTGTGTGATGTGACTAATTTGATAGCAT	812
Db	715	CACCAT	GGACCATTTATTTGTCATACCTGTGTGATGTGACTAATTTGATAGCAT	794
OY	813	CAAAAG	ATTCGGGGGCAACAGCTCGAACAATTTATTCGCAACATGGGCTCATAGAAG	872
Db	795	CAAAAG	ATTCGGGGGCAACAGCTCGAACAATTTATTCGCAACATGGGCTCATAGAAG	854
OY	873	CATCTT	GGAGATTTTAATAAAGACAGATATCAAAATTCCTGAGGACTGGCCTTACCAGA	932
Db	855	CATCTT	GGAGATTTTAATAAAGACAGATATCAAAATTCCTGAGGACTGGCCTTACCAGA	914
OY	933	AGCTGCA	GCCTTGTTCGAAGGACCTATATGCACCTTGGGATATTCCTGAGCTTAAATGGAC	992
Db	915	AGCTGCA	GCCTTGTTCGAAGGACCTATATGCACATTTGGATATTCCTGAGCTTAAATGGAC	974
OY	993	TGCACCT	GATGAGGAGGTCTCATATAGTTTCTGGTAAAGATATGTTTCATGAGA	1052
Db	975	TGCACCT	GATGAGGAGGTCTCATATAGTTTCTGGTAAAGATATGTTTCATGAGA	1034
OY	1033	TGCGGT	TGACAAAGCCATAGAAAGATCAATCTGCCAAGATTAATCGTCCCAAGGAA	1112
Db	1033	TGCGGT	TGACAAAGCCATATAGGAATCAATCTGCCAAGATTAATCTGTCCCAAGGAA	1094



RESULT 4
AAAZ7925
ID    AAAZ7925 standard; cDNA; 1381 BP.
XX
AC    AAAZ7925;
XX
DT    12-SEP-2000 (first entry)
XX
DE    Maize Rad2/FEN-1 cDNA.
XX
KW    Maize; Rad2/FEN-1; transgenic plant; male sterile plant; KM    endonuclease; exonuclease; DNA repair; gene targeting; ss. XX
OS    Zea mays.
XX
FH    Key                 Location/Qualifiers FT    CDS                 37..1176 PT    PT                 /*tag= a
PN    WO200036109-A1.
XX
PD    22-JUN-2000.
XX
PF    16-NOV-1999;      99WO-US27147.
XX
PR    15-DEC-1998;      98US-O112332.
XX
PA    (PION-) PIONEER HI-BRED INT INC.
PI    Mahajan PB;
XX
DR    WPI; 2000-452026/39.
DR    P-PADB; AAY95309.
XX
PT    Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA PT    recombination and repair in transgenic plants, e.g. for gene targeting XX    and the production of male sterile plants - XX
PS    Example 1; Page 75-76; 85pp; English.
XX
CC    The present sequence is that of maize cDNA coding for RAD2/FEN-1 CC    (see AY95309). The cDNA was isolated from a cDNA library prepared CC    from W23 line tassel (16-18 cm long) polyA RNA. Rad2/FEN-1 is a CC    structure specific endonuclease which under certain conditions also CC    acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to CC    produce the Rad2/FEN-1 polypeptides in transgenic plant cells. CC    The protein is involved in the regulation of DNA repair and CC    recombination in plant systems and therefore may be used for

Db 781 AAACCTATTGCTCACATGGGTCATAGAAAGCATCTTGAGATCTTAATAAGACAGA 840

QY 901 TATCAATTCCTGAGAGCTGGCTTACCAAGAGCTGAGCGCTGTTTCAAGAGCCTAAT 960  
DB 841 TATCAATTCCTGAGAGCTGGCTTACCAAGAGCTGAGCGCTGTTTCAAGAGCCTAAT 900  
QY 961 GTCACATTCGATATTCCTGAGCTGAGAAATGAGCTGACCTGATGAGAGGCTCATTAAGT 1020  
DB 901 GTCACATTCGATATTCCTGAGCTGAGAAATGAGCTGACCTGATGAGAGGCTCATTAAGT 960  
QY 1021 TTCCTGTAAGATATATGTTTCATGAAATGCGGTGACCAAGGCCCTAAGAAATATC 1080  
DB 961 TTCCTGTAAGATATATGTTTCATGAAATGCGGTGACCAAGGCCCTAAGAAATATC 1020  
QY 1081 AAATCTGCAAGATATATGTTTCGCAAGAGAGCTGAGTCTTTTCAAGCCTAAGTCC 1140  
DB 1021 AAATCTGCAAGATATATGTTTCGCAAGAGAGCTGAGTCTTTTCAAGCCTAAGTCC 1080  
QY 1141 ACCACATTCAGCAGCGCTGAGAAAGAGAGAGCTTCGATTAACAAAGAGAGCTGCG 1200  
DB 1081 ACCACATTCAGCAGCGCTGAGAAAGAGAGAGCTTCGATTAACAAAGAGAGCTGCG 1140  
QY 1201 AACAGAGAGAGAGAGCTGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1141 AACAGAGAGAGAGAGCTGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
QY 1261 CGACTACGAAAGAGAGCGGTGATGATCACTTGGCTGATTAATTAATCTCCCTGTTTAA 1320  
DB 1201 CGACTACGAAAGAGAGCGGTGATGATCACTTGGCTGATTAATTAATCTCCCTGTTTAA 1260  
QY 1321 ACTAGAGCTTTGGTGAAGATTTGCCATGTTTCAAGCTGGGTAAATAGTTGTTGTTTG 1380  
DB 1261 ACTAGAGCTTTGGTGAAGATTTGCCATGTTTCAAGCTGGGTAAATAGTTGTTGTTTG 1320  
QY 1381 AAGAGATTGGTGTACCAAGTAACAAACTTATCGCTGTTT 1423  
DB 1321 AAGAGATTGGTGTACCAAGTAACAAACTTATCGCTGTTT 1383

RESULT 5  
AA02117  
ID AA02117 standard; DNA; 1144 BP.

AC AA02117;  
DT 23-APR-1999 (first entry)  
XX  
DE Human FEN-1 DNA fragment.  
XX  
KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;  
KW neoplasma; antineoplastic agent; cleavage; ss.  
XX  
OS Homo sapiens.  
XX  
PN US874283-A.  
XX  
PD 23-FEB-1999.  
XX  
PF 30-MAY-1995; 95US-0455968.  
XX  
PR 30-MAY-1995; 95US-0455968.  
XX  
PA (HARR/) HARRINGTON J J.  
PA (HSIE/) HSIEH C.  
PA (LIEB/) LIEBER M R.  
XX  
PI Harrington JJ, Hsieh C, Lieber MR:  
XX  
DR WPI: 1999-179985/15.  
XX  
PT DNA encoding flap endonuclease polypeptides - useful for producing  
XX e.g. recombinant polypeptides  
XX  
PS Claim 3; Column 25; 58pp; English.  
XX

CC This sequence encodes a human FEN-1 (flap endonuclease) protein. This  
CC protein can be used in methods for detecting a pathological condition in  
CC a patient, for diagnostic purposes, for screening for antineoplastic  
CC agents and carcinogens, for diagnostic staging of neoplasia, for  
CC producing recombinant flap endonuclease for use as research or  
CC diagnostic reagents, for producing antibodies reactive with the novel  
CC polypeptides, for producing transgenic nonhuman animals expressing the  
CC novel polypeptides encoded by a transgene. The invention also provides  
CC novel molecular cloning techniques and reagents involving cleavage of  
CC a flap or nick with a flap endonuclease.  
XX

Sequence 1144 BP; 292 A; 283 C; 353 G; 216 T; 0 other:

Query Match 20.3%; Score 299.4; DB 20; Length 1144;

Best Local Similarity 56.5%; Pred. No. 4.3e-70;

Matches 599; Conservative 0; Mismatches 456; Indels 6; Gaps 2;

QY 97 ATGGCATCAAGAGGTTTACGAAACTGCTGGCGGACAAATGCGCCGACGATGAGCATC 156  
DB 1 ATGGCAATTCAGAGCGCTGCGCAAACTAATTCGTATGTGGCCGCCAGTCCATCGGAG 60  
QY 157 CAGAGTTCGAGAGCTACTTGGCGCGCAAAATCGCCGACGATGAGCATC 216  
DB 61 AATGACATCAAGAGCTACTTGGCGCGTAAAGTGGCATGATGCTCTATAGCATTTAT 120  
QY 217 CAGTTCCTGATAGTATGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 276  
DB 121 CAGTTCCTGATGCTGTTGCGCAG---GGTGGGATGTGTGCAAGAAATGAGAGAGGTG 177  
QY 277 GTCACTAGCATTTGCAAGAGAAATGTTCAACCGGACAAATTAATTAATTAATTAAT 336  
DB 178 ACCACAG 237  
QY 337 AAGCCAGTTATGTTTGTGATGAGCAAGCCCTCGTATTAAGAAACAAAGAACTGTTAA 396  
DB 238 AAGCCAGTTATGTTTGTGATGAGCAAGCCCTCGTATTAAGAAACAAAGAACTGTTAA 297  
QY 397 AGATACCAAAAGAGATGAGCAACCAAGATCTGACTGAGGAGAGAGAGAGAGAGAT 456  
DB 298 CCGAGTGAAGCGGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 357  
QY 457 AAGAGATGATGAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516  
DB 358 GAGCAGAT 417  
QY 517 GATTGAAACGACATTAAGACTTATGAGGAGTTCCTGTTGAGAGAGAGAGAGAGAG 576  
DB 418 GAGTGAAG 477  
QY 577 GAGCAGAT 636  
DB 478 GAGCCAG 537  
QY 637 AAGGATCCCTTACTTTGGGGCTCCAGAGTCCCTTGGCATTTATTAATTAATTAAT 696  
DB 538 ATGAGCTGCTCACTTGGCGAGCCCTGCTGTAATGCAACACCTGAGCTGCCAGAGCC 597  
QY 697 AAGAAATACCTGTATGATGAATTTGATGTTGCCAAGTTTGGAGAGAGCTGAATCACC 756  
DB 598 AAAAGAGCTCAATCCAGAGAAATTCACAGCTGAGCGGATTTGACAGAGCTGGGCTGAG 657  
QY 757 ATGAGCAGTTGATGATTTGTCATCCTGTGTGATGATGATGATGATGATGATGAT 816  
DB 658 CAGGAACAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717  
QY 817 GGTATCGGGGGGCAAGAT 876  
DB 718 GGTATCGGGGGGCAAGAT 777  
QY 877 TTGAGAGATCTTAATTAACAGATATCAATTCCTGAGAGAGAGAGAGAGAGAGAGAT 936  
DB 778 GTGCGGAT 837

Accession	Gene	Protein	Location/Qualifiers
QY 937	CGAGCCTTGTCANAGACCCTAAAG---	TCACCTTGGATATCTCTAGGCTAAATGCACT	993
Db 838	CACCAAGCTCTTCTTGGAAACCTGAGGTGCTGGACCCAAAGTCTGTGAGAGCTGAAGTGGAGC		897
OY 994	GCACCTGATGAGAGAGGGCTCATAGTTTCTGTGTAAGAATATGTTTCATGAAGAT		1053
OY 1054	CGGGTGACAAAGGCCATAGAGAGATCAAAATCTGCCAAGATAAATGTCGCCAAGAGAGA		1113
Db 958	CGAATCCGACAGTGGGGTCAAGAGGCTGAGTAAAGACGCCAAGCACCACCCAGGGCGCC		1017
OY 1114	CTCGAGTCCCTTTTTCAGGCCCACTGCCACCATCAGCACC		1154
Db 1018	CTGGATGATTTCTTCAAGGTGACGGGCTCACTCTTTCAGC		1058
RESULT 6			
AAK02107	AAK02107	standard; cDNA; 1144 BP.	
XX AC	AAK02107;		
XX DT	23-APR-1999	(first entry)	
XX DE	Human FEN-1 cDNA.		
XX KW	FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;		
XX KM	neoplasm; antineoplastic agent; cleavage; ss.		
XX OS	Homo sapiens.		
XX FT	Key	Location/Qualifiers	
XX FT	CDS	1..1144	
XX FT		/*tag= a	
XX FT		/product= "FEN-1"	
XX PN	US5874283-A.		
XX PD	23-FEB-1999.		
XX PE	30-MAY-1995; 95US-0455968.		
XX PR	30-MAY-1995; 95US-0455968.		
XX PA	(HARR/) HARRINGTON J J.		
XX PA	(HSIEH/) HSIEH C.		
XX PA	(LIEB/) LIEBER M R.		
XX PI	Harrington JJ, Hsieh C, Lieber MR;		
XX DR	WPI: 1999-179985/15.		
XX DR	P-PSDB: AAM92504.		
XX PT	DNA encoding flap endonuclease polypeptides - useful for producing		
XX PT	e.g. recombinant polypeptides		
XX PS	Disclosure; Fig 1B; 58pp; English.		
XX CC	This sequence encodes a human FEN-1 (flap endonuclease) protein. This		
XX CC	protein can be used in methods for detecting a pathological condition in		
XX CC	a patient, for diagnostic purposes, for screening for antineoplastic		
XX CC	agents and carcinogens, for diagnostic staging of neoplasia, for		
XX CC	producing recombinant flap endonuclease for use as research or		
XX CC	diagnostic reagents, for producing antibodies reactive with the novel		
XX CC	polypeptides, for producing transgenic nonhuman animals expressing the		
XX CC	novel polypeptides encoded by a transgene. The invention also provides		
XX CC	novel molecular cloning techniques and reagents involving cleavage of		
XX CC	a flap or nick with a flap endonuclease.		
XX SO	Sequence 1144 BP; 292 A; 283 C; 353 G; 216 T; 0 other;		

	Best Local Matches 599;	Similarity 56.5%;	Pred. No. 4.3e-70;	Conservative 0;	Mismatches 456;	Indels 6;	Gaps 2;
QY	97	ATGGGCATCAAGAGGTTTGACGAACCTGCTGCGGACAAATGCCCCCAAGCGCATGAAGAG	156				
Db	1	ATGGCAATTCAMGGCCTGGCCAAACTAATTCGTGATGTGGCCCGCAGTCCGCGGAG	60				
QY	157	CAGAACTTGAAGACTACTCTGGCCGCAAAATGCCGCTGACGCCACGATGACATCTAC	216				
Db	61	AATGACATCMAAGCTACTTTGGCGTAAAGGGCCATTTGAATGCCCTTAATGAGCATTTAT	120				
QY	217	CAGTTCCTCATGTAGTTGTAAGGACGAGCATGGAACCTTCACAATGAATCAACTGGTGA	276				
Db	121	CAGTTCCTCATGTAGTTGTAAGGACGAGCATGGAACCTTCACAATGAATCAACTGGTGA	177				
QY	277	GTCACCTAGTCATTTGCCAAGGAATGTTCAACCGGACAAATTAAGATTACTGGAACGGGAATC	336				
Db	178	ACCACCGACGACCTCATGGGACATGTTCACCCACCATTCGCATGTAGGAACGGCATC	237				
QY	337	AAGCCAGTTTATGTTTTTATGGAAGCAAGCCTCCTGATATGAACAACAGAACTTCTTAA	396				
Db	238	AAGCCCGTATATCTTTGATGGAAGCCGCAACGCTCAATCTACGGCAGCTGGCCAAA	297				
QY	397	AGATACCTCAAAAGAGATGATGACACCAAGATCTGACTGAGGACGATAGAGTACAGAT	456				
Db	298	CCGACTGACGGCGGGCTGAGGCAAGAAAGACGTGCAACAGGCTCAGGCTCCTGGGCGC	357				
QY	457	AAAGATGCCATTGAAAATTTGACGAAGAGACTGTAAAGTACACAGGACACAAACGA	516				
Db	358	GAGCAGGAGAGTGGAAAAATTCATAAGCGCGTGGGAAGGTCAATGAAGCAGCATGAT	417				
QY	517	CATTCTAAACGACTTTTAAGACTTTAGGGGGGTTCCTGTTGTAGAGCAGCTTCTTACGA	576				
Db	418	GAGTCAAAACATCTGCTGAGCCTCATGGGCAATCCCTTACTTGATGACACCCAGTAGGCA	477				
QY	577	GAGCAGAAATGTGACCCCTTTGCATTAACGATTAAGTGTTGGCTTGCTTACAGACAT	636				
Db	478	GAGCGCAGCTGTGCGCCCTGGTGAAGGCTGCAAAAGTCTATGCGGCTACCGAGAC	537				
QY	637	AAGGACCTCCTACTTTTGGGGCTCCACGCTTCCTTCGTCATTATAGATCCAGTTCC	696				
Db	538	ATGGACTGCTTACCTTCGCGACCCCTGTGCTAAATGCGACACCTCACTGCGAGAACG	597				
QY	697	AAGAAATACCTGTCATGGAATTTGATGTGTCGAAGGTTTTGGAGAGCTTGAATCAC	756				
Db	598	AAAAAGTCCCAATCCAGGAATTCACCTGACCGCGGATTTCTCAGAGAGCTGGCGCTGAC	657				
QY	757	ATGGACCACTTATTTGATTTGTGCATCTCTGTGTCATGTGACTATTGTGATACATCAA	816				
Db	658	CAGGCAACATTTGTGATCTGTGCACTCTGCTGAGGACGAGTACTGTGAGAGTATCCGG	717				
QY	817	GATATCGGGGGGCAAAACGCTCTGAACCTTTATTCGCAACATGGGTCATGAAAGCATC	876				
Db	718	GGTATTTGGGCCCAACGGGCTGTGCACTCATCCAGAAACACAAAGCATTCAGGAGATC	777				
QY	877	TTGGAATCTTAAATTAAGACAGATATCAAAATTCCTGAGAGCTGGCCTTACCAAGACT	936				
Db	778	GTCGGCGCACTTGACCCCAACAAAGTAACTCTGGCCAGAAATTTGCTCCACAAAGAGCT	837				
QY	937	CGACCTGTTTCAAGAGGCTTAAGT---TCACTTTGGAATTTCCAGACTTAAATGACT	993				
Db	838	CACCAAGCTCTTTTGTGAACCTTGAGTGTCTGGACCCAGAGTCTGTGACCTGAAAGTGAGC	897				
QY	994	GCACCTGATGAAGAGGATCTCATTAAGTTTCCGTGTAAGAAAGATTAATGTTTCAATGAAT	1053				
Db	898	GAGCCAATTAAGAAAGAGCTGTATCAATTTTCATGTGTGTAAGAAACAGTCTTCAGAG	957				
QY	1054	CGGGTGACAAAGCCCATAGAGAAGTCAAAATCTGCCAAGATTAATTCGTGCAAGGAAGA	1113				
Db	958	CGAATCCGACATGAGGGTCAAGAGGCTGTAGTAAGACCCGCCCAAGGACGACCAACCGAGGCCG	1017				
QY	1114	CTCGAGTCTTTTTCAGCCCAACTCCCAACCATATGAGACC	1154				

Db 1018 CTGATGATTTCTTCAAGTGACCGGCTCACTCTTTCAGC 1058

RESULT 7  
AA02111  
ID AAX02111 standard; DNA; 2033 BP.  
XX  
AC AAX02111;  
XX  
DT 23-APR-1999 (first entry)  
XX  
DE Human FEN-1 genomic DNA.  
XX  
KM FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;  
KW neoplasm; antineoplastic agent; cleavage; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 104..1240  
FT /tag= a  
FT /product= "FEN-1"

XX US5874283-A.  
XX  
XX 23-FEB-1999.  
XX  
XX 30-MAY-1995; 95US-0455968.  
XX  
XX 30-MAY-1995; 95US-0455968.  
XX  
XX 30-MAY-1995; 95US-0455968.  
XX  
XX (HARR/) HARRINGTON J J.  
XX (HSIEH/) HSIEH C.  
XX (LIEBER/) LIEBER M R.  
XX  
XX Harrington J, Hsieh C, Lieber MR;  
XX  
XX WPI: 1999-179985/15.  
XX P-PSDB: AAM92508.  
XX  
XX DNA encoding flap endonuclease polypeptides - useful for producing  
XX e.g. recombinant polypeptides  
XX  
XX PS Disclosure: Fig 5A-B; 58pp; English.  
XX  
XX This sequence encodes a human FEN-1 (flap endonuclease) protein. This  
XX protein can be used in methods for detecting a pathological condition in  
XX a patient, for diagnostic purposes, for screening for antineoplastic  
XX agents and carcinogens, for diagnostic staging of neoplasia, for  
XX producing recombinant flap endonuclease for use as research or  
XX diagnostic reagents, for producing antibodies reactive with the novel  
XX polypeptides, for producing transgenic nonhuman animals expressing the  
XX novel polypeptides encoded by a transgene. The invention also provides  
XX a novel molecular cloning techniques and reagents involving cleavage of  
XX a flap or nick with a flap endonuclease.  
XX  
XX Sequence 2033 BP; 500 A; 493 C; 579 G; 461 T; 0 other;

Query Match 18.6%; Score 275; DB 20; Length 2033;  
Best Local Similarity 55.8%; Pred. No. 2e-63;  
Matches 637; Conservative 0; Mismatches 480; Indels 24; Gaps 5;

Db 97 ATGGGATCAAGGAGTTCAGCAAACTGCTGGCGGACATGCGCCCAAGCGATGAAGAG 156  
1111 11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 104 ATGGAATTCACGGCTTGCCAACTAATTGCTGATGTGGCCGCCAGTGCCATCCGTAG 163  
1111 11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
QY 157 CAGAAGTCGAGAGCTACTTGGCCGCAAAATCGCCGTCGACGCGCATGAGATCTAC 216  
11 11 11111111 11 1111 11 1111 11 11111111 1111111111  
Db 164 AATGACATCAAGAGCTATTGTGTCCCAAGTGGCGCATGCTCCTCATAGCATCTAC 223  
1111111111 1111 11 1111 11 1111 11 1111 11111111  
QY 217 CAGTTCCGATAGTAGTTGGAGGACGATGCAAACTCTCACAAATGAAGCTGGTGA 276  
1111111111 11 11 11 1111 11 1111 11 1111 11111111  
Db 224 CAGTTCGATTTGCTTCTGTAC---GGTGGGATGTGTGACGAAGAGAGGAGTGA 280  
1111111111 11 11 11 1111 11 1111 11 1111 11111111

QY 277 GTCACTAGTCATTTGCAAGGAATGTTCAACCGGACAAATAGATTACTGGAAGCGGAATC 336  
1111 11 11 11111111 11 11 11 1111 11 1111 11 1111  
Db 281 ACCACAGACC---TGATGGCATGTGTTACCTGACCATGCG---CATGGAAATGGCATC 334  
1111 11 11 11111111 11 11 11 1111 11 1111 11 1111  
QY 337 AAGCCAGTTTATGTTTTTGATGGCAAGCCTCTGATATGAAAGAAACAAGAACTGGCTAA 396  
1111 11 11 11111111 11 1111 11 1111 11 1111 11 1111  
Db 335 AAGCCTGTGTGCTTGTGATGGCAAAACCCACAGAGTCAAGAGTCAAGGAGCTGGCCAA 394  
1111 11 11 11111111 11 1111 11 1111 11 1111 11 1111  
QY 397 AGATACATCAAAAAGATGATGCAACCAAGATCTGACTGAGGCAATGAGTGGAGAT 456  
11 11 11 1111 11 11 11 1111 11 1111 11 1111  
Db 395 CGCAGTGAGAGCGCGCCGAGGCTGAGAGCACTGACGAGCTCAGCAGGCTGGGATG 454  
11 11 11 1111 11 11 11 1111 11 1111 11 1111  
QY 457 AAATGCGATTTGAAAAATTGACGAGAGCTGTAAGACGACCAAGCAACAGCAAGCA 516  
11 11 11 11111111 11 1111 11 11111111 11111111 11  
Db 455 GAGAGGAGGTGGAGAACTTACCAGAGGCTGCTGAAGGTCCACCAAGCAACATGAT 514  
11 11 11 11111111 11 1111 11 1111 11 1111 11 1111  
QY 517 GATTGTAAACGACATTAAGACTTATGAGGAGGCTGCTGTGAGAGGACCTTGAAGCA 576  
11 11 11 1111 11 11 11 1111 11 1111 11 1111  
Db 515 GAGTGCAAACACTGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 574  
11 11 11 1111 11 11 11 1111 11 1111 11 1111  
QY 577 GAACAGAGATGTGAGGCTTGGATTAAGCATAGGTGTGCTGCTTCCAGAGAT 636  
11 11 11 1111 1111 11 11 11 1111 11 11 11  
Db 575 GAGGCAAGCTGTGCTGCTGCTGCAAGAGCTGCAAGAGCTATGCTGCGGCAAGAGAG 634  
11 11 11 1111 1111 11 11 11 1111 11 11 11  
QY 637 AAGGACCTCCCTTACTTTTGGGCTCCACGCTTCCCTGCTGCTGCTGCTGCTGCTGCT 696  
111111 1111111111 11 11 11 1111 11 11 11 1111 11  
Db 635 ATGACCTGCTCACTTTTGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 694  
111111 1111111111 11 11 11 1111 11 11 11 1111 11  
QY 697 AAGAAATACCTGTGATGGAATTTGATGTTGCAAGGTTTGGAGAGCTGGAATCC 756  
1111 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 695 AAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 754  
1111 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
QY 757 ATGACCTGCTCACTTTTGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 755 CAGGAGAGTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
QY 817 GGTATCGGGGGGCAACACTCTGAACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 815 GGCATTTGGCGCCCAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 874  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
QY 877 TTGGAATCTTAATTAAGACAGATATCAAAATCTGAGAGCTGCTGCTGCTGCTGCTGCT 936  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 875 GTGAGGCGGCTGAGACCCAGCAAGTACCCGCTGCAAGAGTACGCTGCTGCTGCTGCTGCT 934  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
QY 937 CGAGCTTGTTCAGGAGCCCTAA---TGTCACTTTGATTTCTGAGCTTAAATGAGACT 993  
11 11 11 1111 1111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 935 CAGAGCTTCTTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 994  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
QY 994 GCACCTGATGAGGAGGCTGCTGATGATTTCTGTAAGATTAATGTTCAATGAAGAT 1053  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 995 GAGCAATTAAGAAAGATTTGCTCAAAATTTATGTGTGTAAGAAAGCAATTTTGTGAAGAG 1054  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
QY 1054 CGGGTACAAAGGCTGATGAGAGATCAAAATCTGCAAGAAATTAATCTGCAAGAGAA 1113  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 1055 CGAATTCGAGTGGGCTGCAACCGGCTGAGTAAGAGCCCGCAGGAGCAGCAGCAGGAGAGC 1114  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
QY 1114 CTCGATCTTTTCAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1173  
1111 1111 11111111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 1115 CTCGATGATTTCTTCAAGTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152  
1111 1111 11111111 11 11 11 1111 11 11 11 1111 11 11 11  
QY 1174 TCGGATTAAGCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1233  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 1163 AAGGAGCCAGAAACCAAGGAGGCTGCTGTAAGAAAGAAAGACTGCGGAGCGGAGAG 1222  
11 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
QY 1234 T 1234  
1111 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11  
Db 1223 T 1223  
1111 11 11 1111 11 11 11 1111 11 11 11 1111 11 11 11

RESULT 8  
ABL14287  
ID ABL14287 standard; cDNA; 1300 BP.



```

PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
PI Harrington JJ, Hsieh C, Lieber MR;
XX
DR WPI; 1999-179985/15.
DR P-PSDB; AAM92505.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing
PS e.g. recombinant polypeptides
XX
PS Disclosure; Fig 2B; 58bp; English.
XX
CC This sequence encodes a mouse FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel
CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.
XX
SQ Sequence 1930 BP; 482 A; 459 C; 553 G; 436 T; 0 other:

Query Match 18.3%; Score 270.6; DB 20; Length 1930;
Best Local Similarity 55.0%; Pred. No. 2,9e-62;
Matches 628; Conservative 0; Mismatches 489; Indels 24; Gaps 4;

OY 97 ATGGGATTCAGAGGCTTGGGCAAACTAATGCTGATGTGGCCCGCCAGTCCGCTGAG 60
DB 1 ATGGGATTCAGAGGCTTGGGCAAACTAATGCTGATGTGGCCCGCCAGTCCGCTGAG 60
OY 157 CAGAGTTGAGAGTACTTCCGCGCAAAATGCGCGAGCGAGCATGAGCATCTAC 216
DB 61 AATGACATCAAGAGCTACTTGTGCTGTAAGTGGCCATGATGCCATGAGCATCTAC 120
OY 217 CAGTTCTGTAGTAGTGTGAGAGAGCATGAAACTCTCAAAATGAAGCTGTGAA 276
DB 121 CAGTTCTGTAGTAGTGTGAGAGAGCATGAAACTCTCAAAATGAAGCTGTGAA 177
OY 277 GTACTGTGATTTGCAAGAAATGTTCAACCGAGCAATAAGATTACTGAGAGCGGATC 336
DB 178 ACCACGAGCTGATGGGATGTTAT-----GGCAACCATCCGCATGGAGATGCGATC 231
OY 337 AAGCAGTTTATGTTTGTATGGCAAGCTCTGATATGAAGAAAGCAAACTTGTATAA 396
DB 232 AAGCCTGTGATGCTTGTATGGCAAAACCAACAGCTGAAGTGAAGCGAGTGGCCAG 291
OY 397 AGATATCAAAAAGATGATGCAACCAAGATCTGACTAGGCAAGTAGAGTAGAT 456
DB 292 CGCATGAGAGGCGCGCGAGGCTGAGAGAACCACTGACAGGCTCAGAGGCTGGAGT 351
OY 457 AAAAGTGGATTTGAAAATTTGAGCAAGAGAGCTGAAGTCAAGCAAGCAACAGAA 516
DB 352 GAGGAGAGAGGTGGAGAAATTCACCAAGAGGCTGCTGAAGTCAACAGCAACAAATGAT 411
OY 517 GATTGTAAGAGCATTTAAGACTTATGAGGGGTTCTGTTTGAAGGCACTTGTGAACA 576
DB 412 GAGTGCAAAACCTCGTAGAGCTCATGGGATCCCTTACCTTGATGACACCAGAGGCA 471
OY 577 GAAGCAGATGTGAGCCCTTTGCTATAAGATAGTGTTCGCTGTTGTTTGAAGAAT 636
DB 472 GAGGCCAGCTGTGCTGCTGCGCAAGGCTGCGCAAAAGTCTAGTGTGGGCGAGGAG 531
OY 637 AAGGACTCCCTTACTTTTGGGGCTCCACGGTTCCTTGCTGATTTAAGGATCCAAAGTTC 696
DB 532 ATGGACTGCTCCTACTTTTGGCAGCCCGCTGCTAATGGACACCTAATCTGCAGTGAAGCC 591

```

```

OY 697 AAGAAATACCTGTGATGAGATTTGATGTTGCCAGGTTTGGAGAGCTTGAACCTACC 756
DB 592 AAGAAGCTGCCATCCAGAGGTTCTGAGCGCGCTCTGTGAGAGAGCTGGGTGGAAC 651
OY 757 ATGAGCAGCTTCATTTGATTTGTCATCTGTTGTGATGTGACTATTGTATGATCAAA 816
DB 652 CAGGAGCAGTTTGTGATCTGTGATCTGCTGAGTGTGAGTACTAGTCAAGAGATCCCT 711
OY 817 GGTATCGGGGGGCAAGACGCTGGAACCTTATTCGTCAACATGGGCTCCATAGAAGCATC 876
DB 712 GGCATTTGGCGCCAAAGCGGCTGTGATCTCATCCAGAAACATAGAAGCATCGAGGATTC 771
OY 877 TTGAGATCTTTAATAAGACAGATATCAATTCCTGAGGACTGGCTTACCAAGAAGCT 936
DB 772 GTGAGCGGGGTGAGACCCCAAGTACCCGTTCCAGAGAACTGGCTCCAGAGAGGCC 831
OY 937 CGAGCGTTTGTCAAGAGCCTAA---TGTACTTTTGATATTCTCTGAGCTAAATGAGCT 993
DB 832 CAGCAGCTTCTTCGAGGAGCCAGAGTAGTGACCACAGTCTGTGAGCTGAAGTGGAGC 891
OY 994 GCACCTGATGAGGAGGCTCATATAGTTTCTGTTAAAGATATGATGTTCAATGAAGAT 1053
DB 892 GAGCCAAATGAAGAAGATTGTTAAATTTATGTGTGTGAAAAAGCAGTTTCTGAAGAG 951
OY 1054 CGGCTGACAAAGGCCATAGAGAGATCAATCTGCCAAGATTAATTCGCCAAGAGAGA 1113
DB 952 CGAATTCGCAATGGGGTCAACGCCGCTGATAGAGAGCCGCCAGGAGCAGACCCAGGAGCG 1011
OY 1114 CTCGAGTCTTTTTCAGGCCACTGCGACCATACACACCGCGTAAAGAGAGAGACT 1173
DB 1012 CTCATGATTTTCTTCAAGGTGACAGGCTCACTCTCTCAGC-----TTAGCCG 1059
OY 1174 TCGGTAATAACAGACAGGAGCTGCGACAGAAAGAAAGAGCTGTGGAAGAAAGAAA 1233
DB 1060 AAGGAGCCAGAGACCAAGGGGCTGCTAAGAAAGAAAGCAAGATGGGAGAGGGGAG 1119
OY 1234 T 1234
DB 1120 T 1120

RESULT 10
AAK02109
ID AAK02109 standard; cDNA; 1149 BP.
XX
AC AAK02109;
XX
DT 23-APR-1999 (first entry)
XX
DE Yeast FEN-1 cDNA.
XX
KW FEN-1; yeast; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasia; antineoplastic agent; cleavage; ss.
XX
OS Saccharomyces cerevisiae.
XX
PN US5874283-A.
XX
PD 23-FEB-1999.
XX
PF 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
PI Harrington JJ, Hsieh C, Lieber MR;
XX
DR WPI; 1999-179985/15.
DR P-PSDB; AAM92505.
XX

```









```

Db 244 GAGGACCAAGATGCCCGTCAAGAGATTACGTAACGAAAGCTGTGGAGGCTGGCC 303
QY 751 CTCACCATGAGCAAGTTCATGATTTGCAATCTGTGTGATGTGATATGTGATAGC 810
Db 304 ATTACCAATGCGAGTTCATGATGATATGCTGCTGCGGTTGGGATTTACTGTAGAC 363
QY 811 ATCAAGATATCGGGGGGCAACAGCTCTGAACCTTATTCGTCAACATGGGTTCATAGAA 870
Db 364 ATCAAGGGTATTTGGACCCAGCGATCGAACTGATCAACACCTATCGGATATAGAG 423
QY 871 AGCATCTTGGAATCTTATTAAGACAGATATCAATTCCTGAGACACTGGCTTACCAA 930
Db 424 ACTATTTCTGGATTAACCTGACTGACTAGCAAAATACACCTGCGGAGAACTGGAAC 483
QY 931 GAACCTCGACGCTTTGTCAGAGAGGCTTATGCTCACTTTGATATTCCT---GAGCTAAA 987
Db 484 GTGGCGGGGAAACTCTTCATGAAACCGAGGTAGCTGATGCCGACTCCATAGATTTCAA 543
QY 988 TGGACTGACCTGATGAGAGAGGCTCTCAVAAAGTTTCTGTGTAAGATATATGTTTCAAT 1047
Db 544 TGGGTGAGCGGATGAGAGAGGCTTGTCAAGTTTCTGTGCGGAGCCGCGAGTTCAAC 603
QY 1048 GAAGATCGGGTGACAAAGGCCATAGAGAAGATCAATCTGCCAAGATTAATTCGTGCAA 1107
Db 604 GAAGAGCCCGTTCCGACAGCGTCCCAAAAAGCTGATGAATCCAAAGCCAGCCAGCTCAG 663
QY 1108 GGAAGACTCGAGTCTTTTTCAGGCCCAACTGCCACAC 1145
Db 664 GTGAGACTCGATAGCTTCTTTAAGACACTGCCACGAC 701

RESULT 13
AAT76685 standard; DNA; 1023 BP.
AC AAT76685;
AT 14-APR-1998 (first entry)
DE Pyrococcus furiosus FEN-1 endonuclease gene coding sequence.
XX
KM Nucleic acid cleavage; 5' nuclease; DNA cleavage; RNA cleavage;
KM Invader directed cleavage; FEN-1; endonuclease; ds.
XX
OS Pyrococcus furiosus.
PN W09727214-A1.
PD 31-JUL-1997.
XX
PF 22-JAN-1997; 97MO-US01072.
PR 02-DEC-1996; 96US-0759038.
PR 24-JAN-1996; 96US-059491.
PR 12-JUL-1996; 96US-0682853.
PR 29-NOV-1996; 96US-0756386.
PR 02-DEC-1996; 96US-0758314.
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichiev VI;
PI Olive DM, Prudent JR.
XX
DR P-PSDB; AAW24216.
XX
PT Thermostable structure-specific nuclease(s) - used for detection and
PT characterisation of nucleic acid sequences and variations in nucleic
XX acid sequences
XX
PS Example 28; Page 283-285; 457pp; English.
XX
CC This sequence comprises the coding region of the gene encoding

```

```

CC Pyrococcus furiosus (Pfu) FEN-1 endonuclease (see AAW24216).
CC It was obtained by PCR amplification (see AAT76682-83). Large
CC scale preparation of recombinant Pfu FEN-1 was performed using
CC E. coli as host. Pfu FEN-1 is a thermostable enzyme. It can be
CC used in novel methods for the detection and characterisation of
CC nucleic acid sequences and variations in nucleic acid sequences.
XX
SQ Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other:

Query Match 11.0%; Score 162.8; DB 18; Length 1023;
Best Local Similarity 55.0%; Pred. NO. 1.8e-33;
Matches 320; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

QY 146 CATGAGAGAGCAGAGTTCCGAGACTACTTTCGGCCGCAAAATTCGCGTCCAGCCAGCA 205
Db 29 CAGAAAGAAATTAAGTTAGTAAACCTATACGAAAAAAATTCGCAATTCGACCTCTTA 88
QY 206 TGAGCATCTACCACTTCCGATAGTAGTGGAGAGACAGCATGGAACCTCCACAATG 265
Db 89 ATGCAATCTACCAATTTTGTCCACATTAAGACAGAAAGATGGAATCTCATTATGAT 148
QY 266 AAGCTGTTGAAGTCACTGTCATTTGCAAGGATGTTCAACCGCAATTAAGATTACTGG 325
Db 149 CAAAGGTAAGATTAACCTCCACCTTAAGCGGCTCTTTTACAGGACAAATTAACCTAATGG 208
QY 326 AAGCGGAATCAAGCCATTTATGTTTGTGATGGCAAGCTCTCTGATATGAAGAAACAG 385
Db 209 AGGCTGGAATTAACCTGTGTATGTTTGTGATGAGAACTTCACAAATTCAAAAAGAAAG 268
QY 386 AACTTGCTAAAGATATCTCAAAAAGAGATGATGCACCAAGATGACTGACGAGTAG 445
Db 269 AGCTCGAAAAAAGAGAGAGAGAGAGAACTGAAAGAAAGATGAGAGAGACACTTG 328
QY 446 AGTATGAGATTAAGATCCGATTGAAAAATTGAGCAGAGACTGTAAAGTTCACAGGC 505
Db 329 AAAAAGGAGATAGAGAGAACCAAAATATGCCCAAGAGCAACAGGCTAATATGAAA 388
QY 506 AACCAACGAGATTTGTAAAGCATATTAAGCATTTAGGGGTTCTCTGTAGAGGAC 565
Db 389 TGCATCATGAGAGATGCAAAAAAAGCTTTAGAGCTTATGGAATTCCTATATGTTCAAGTAC 448
QY 566 CTTCTGAAGCAGAGCAGATGTGCAGCCCTTTGCATTAACGATTAAGGTTCGCTGTTG 625
Db 449 CTACGAGAGGAGAGGCCAAGCTGCATATATGGCCGCAAGGGAGCGGTATGCAATCGG 508
QY 626 CTTCAAGATTAAGACCTCCCTTACTTTTGGGCTCCACGCTTCTTGCTATTAATGG 685
Db 509 CTAGTCAAGATTAGCATTCCTTACTTTTGGAGCTCCAAAGACTTGTAGAAACTTAACAA 568
QY 686 ATCCAGTTCCAGAAATATCACTGTGATGAATTTGATGTTG 727
Db 569 TAACAGGAAAAAGAAAGTTGCTGGAATAATGTCTACGTGG 610

RESULT 14
AAV65840 standard; DNA; 1023 BP.
AC AAV65840;
AT 02-FEB-1999 (first entry)
DE Pyrococcus furiosus FEN-1 endonuclease gene ORF.
XX
KM Nucleic acid detection; multiple sequential invasive cleavage;
KM FEN-1; endonuclease; nuclease; ds.
XX
OS Pyrococcus furiosus.
PN W09842873-A1.
PD 01-OCT-1998.
XX

```









D	301	TTCAACCGGACAAATAGATTACTGGAAGCCGGCAATCAAGCCAGTTTATGTTTTTGATGGC	360
Q	361	AAGCCTCCTGATATGAGAAACAAGAACTTGCTAAAAGATACCAAAAAGAGATGATGA	420
D	361	AAGCCTCCTGATATGAGAAACAAGAACTTGCTAAAAGATACCAAAAAGAGATGATGA	420
Q	421	ACCAAAATCTGACTGAGGCAATGAGGTAGGAGATTAAGATCCGATTGAAAAATTGAGC	480
D	421	ACCAAAATCTGACTGAGGCAATGAGGTAGGAGATTAAGATCCGATTGAAAAATTGAGC	480
Q	481	AAGAGACTGTAAAAGTCACAAAGGCACACAAAGAAATGTGAAGACATTAATAAGCT	540
D	481	AAGAGACTGTAAAAGTCACAAAGGCACACAAAGAAATGTGAAGACATTAATAAGCT	540
Q	541	ATGGGGGTCTCTGTGTGAGAGGCACTTCTGAAGCAGAAACAGAAATGTGAGCCCTTTC	600
D	541	ATGGGGGTCTCTGTGTGAGAGGCACTTCTGAAGCAGAAACAGAAATGTGAGCCCTTTC	600
Q	601	ATAAACGATAGGTGTTCCTGTTGCTTCAGAAAGATTAAGACTCCCTTACTTTGGGGCT	660
D	601	ATAAACGATAGGTGTTCCTGTTGCTTCAGAAAGATTAAGACTCCCTTACTTTGGGGCT	660
Q	661	CCAGGTTCCCTGCTCATTTTAATGGATCCAAAGTCCACAGAAATTAACCTGATGAGAACT	720
D	661	CCAGGTTCCCTGCTCATTTTAATGGATCCAAAGTCCACAGAAATTAACCTGATGAGAACT	720
Q	721	GATGTTGGCAAGGTTTTTGGAGAGCCTTGAACACTACCATGGACCATTGATTTGTGC	780
D	721	GATGTTGGCAAGGTTTTTGGAGAGCCTTGAACACTACCATGGACCATTGATTTGTGC	780
Q	781	ATCCTGTGTGATGTGACTATTGTGATAGCATCAAAAGTATCGGGGGGCAAAAGCTCTG	840
D	781	ATCCTGTGTGATGTGACTATTGTGATAGCATCAAAAGTATCGGGGGGCAAAAGCTCTG	840
Q	841	AAACTTTATGCTCAACATGGGGTCCATGAAGAAGATCTGGAGAACTTAAATAAGCACA	900
D	841	AAACTTTATGCTCAACATGGGGTCCATGAAGAAGATCTGGAGAACTTAAATAAGCACA	900
Q	901	TATCAAAATTCCTGTGAGCACTGGCCTTACCAGAAAGCTGAGGCTTGTTCAGAGAGCCTAAT	960
D	901	TATCAAAATTCCTGTGAGCACTGGCCTTACCAGAAAGCTGAGGCTTGTTCAGAGAGCCTAAT	960
Q	961	GTCACCTTGGATATTTCTGAGCTAAAATGAGACTGCACCTGATGAGAGGGGTCTCATAGT	1020
D	961	GTCACCTTGGATATTTCTGAGCTAAAATGAGACTGCACCTGATGAGAGGGGTCTCATAGT	1020
Q	1021	TTTCCGTGTAAGAAAGTAAATGGTTTCAATGAAGATCGGGTGCAGAAAGCCATAGGAAGATC	1080
D	1021	TTTCCGTGTAAGAAAGTAAATGGTTTCAATGAAGATCGGGTGCAGAAAGCCATAGGAAGATC	1080
Q	1081	AAATCTGCCAAGATTAATATGCTGCCAAGAGAACTCGAGCTCTTTTTCAGGCCAACTGGC	1140
D	1081	AAATCTGCCAAGATTAATATGCTGCCAAGAGAACTCGAGCTCTTTTTCAGGCCAACTGGC	1140
Q	1141	ACCAATCAGCACCGCTTAAAACGAGAGAGACTTGGATTAACCAAGCAAGCAGCTGCG	1200
D	1141	ACCAATCAGCACCGCTTAAAACGAGAGAGACTTGGATTAACCAAGCAAGCAGCTGCG	1200
Q	1201	AACCAAGAAAACAAAGGCTGGGTGGGAAAACAAAGAAATAACTTGGATGCTGATGACAACTA	1260
D	1201	AACCAAGAAAACAAAGGCTGGGTGGGAAAACAAAGAAATAACTTGGATGCTGATGACAACTA	1260
Q	1261	CGACTACGAAAGCAGCGGTGCGATGATCAGCTCGCCTAGATTAATTAACCTCCCTGTTTTA	1320
D	1261	CGACTACGAAAGCAGCGGTGCGATGATCAGCTCGCCTAGATTAATTAACCTCCCTGTTTTA	1320
Q	1321	ACTCAGACCTTTGGTGAAGTTTGGCCATGTTTCAAGCTGGGGTAAAGTTAGTTGTGTTG	1380
D	1321	ACTCAGACCTTTGGTGAAGTTTGGCCATGTTTCAAGCTGGGGTAAAGTTAGTTGTGTTG	1380
Q	1381	AAGAGATTGGTGTACCAAGTAAACAAACTATCGCTTTTTTACTTCTTGTCTTTGAA	1440
D	1381	AAGAGATTGGTGTACCAAGTAAACAAACTATCGCTTTTTTACTTCTTGTCTTTGAA	1440

OY	1441	GTATGTTATGCCAGTAAACCAAAAAAAAAAAAAAAAAAAAAA	1478
Db	1441	GTATGTTATGCCAGTAAACCAAAAAAAAAAAAAAAAAAAAAA	1478
 RESULT 2 US-09-426-557-1			
; Sequence 1, Application US/09426557			
; Patent No. 632527			
; GENERAL INFORMATION:			
; APPLICANT: Mahajan, Pramod B.			
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses			
; TITLE OF INVENTION: Thereof			
; FILE REFERENCE: 0961			
; CURRENT APPLICATION NUMBER: US/09/426,557			
; CURRENT FILING DATE: 1999-10-22			
; EARLIER APPLICATION NUMBER: 60/112,332			
; EARLIER FILING DATE: 1998-12-15			
; NUMBER OF SEQ ID NOS: 10			
; SOFTWARE: fastseq for Windows Version 3.0			
; SEQ ID NO 1			
; LENGTH: 1463			
; TYPE: DNA			
; ORGANISM: zea mays			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (85)...(1221)			
; US-09-426-557-1			
 Query Match                      96.2%; Score 1421.8; DB 4; Length 1463;			
Best Local Similarity        98.5%; Pred. No. 0;			
Matches 1435; Conservative    0; Mismatches    22; Indels     0; Gaps     0;			
OY	19	AATAGCTCGCGGTGGTGCGGTTTCCTGCGCCACTCGCGCTCAGCGCGCCGCCACC GCC	78
Db	7	AATAGCTCGCGCGCGCGGTTTCCTGCGCCACTCGCGCTCAGCGCGCCGCCACC GCC	66
OY	79	ACAGCGCGCGCAGACAGCATGGGCATCAAGGGTTTGAGAAACCTCTGGGGACAATCGG	138
Db	67	ACAGCGCGCGCAGACAGCATGGGCATCAAGGGTTTGAGAAACCTCTGGGGACAATCGG	126
OY	139	CCCCAGGGCGATGAAGAGCAGAGAAGTTCGAGAGCTACTTCGCGCCGCAAAATCGCCCTGCAC	198
Db	127	CCCCAGGGCGATGAAGAGCAGAGAAGTTCGAGAGCTACTTCGCGCGCAAAATCGCCCTGCAC	186
OY	139	GCCAGCATGAGCATCTACCAAGTTCCTGATAGTAGTTGCAAGACAGCGCATGAACCTCTC	258
Db	187	GCCAGCATGAGCATCTACCAAGTTCCTGATAGTAGTTGCAAGACAGCGCATGAACCTCTC	246
OY	259	ACAAATGAAGCGGTGGAGAGTCACTAGTCATTTCGAAGGAATTTCAACCGCACATAAGA	318
Db	247	ACAAATGAAGCGGTGGAGAGTCACTAGTCATTTCGAAGGAATTTCAACCGCACATAAGA	306
OY	319	TTCAGTGAAGCGGGGAATCAAGCCAGTTTATGTTTTGATGGCAAGCTCTTGATTAAG	378
Db	307	TTCAGTGAAGCGGGGAATCAAGCCAGTTTATGTTTTGATGGCAAGCTCTTGATTAAG	366
OY	379	AAACAAGAACCTTGCCTAAAAGATACCTCAAAAGAGATGATGCACCAACCAAGATGTGACTGAG	438
Db	367	AAACAAGAACCTTGCCTAAAAGATACCTCAAAAGAGATGATGCACCAACCAAGATGTGACTGAG	426
OY	439	GCAGTAGAGGTGAGATTAAGATCGGATTTGAAATTTGAGCAAGCAGGACTTAAGGTC	498
Db	427	GCAGTAGAGGTGAGATTAAGATCGGATTTGAAATTTGAGCAAGCAGGACTTAAGGTC	486
OY	499	ACAAGGCAACACACAGAGATTGTAAGGACTATTAGAAGCTTATGGGGTTCTCTTTGA	558
Db	487	ACAAGGCAACACACAGAGATTGTAAGGCGCTATTAGAAGCTTATGGGGTTCTCTTTGA	546
OY	559	GAGGCACTTTCTGAAGCAGAACCAAAATGTGAGCCCTTTGATTAAGCATTAAGGTGTC	618
Db	547	GAGGCACTTTCTGAAGCAGAACCAAAATGTGAGCCCTTTGATTAAGCATTAAGGTGTC	606

```

RESULT 2
US-09-426-557-1
: Sequence 1, Application US/09426557
: Patent No. 6232527
: GENERAL INFORMATION:
: APPLICANT: Mahajan, Pramod B.
: TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
: TITLE OF INVENTION: Thereof
: FILE REFERENCE: 0961
: CURRENT APPLICATION NUMBER: US/09/426,557
: CURRENT FILING DATE: 1998-10-22
: EARLIER APPLICATION NUMBER: 60/112,332
: EARLIER FILING DATE: 1998-12-15
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1463
: TYPE: DNA
: ORGANISM: zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (85) ... (1221)
US-09-426-557-1

```

Query Match	96.2%	Score 1421.8;	DB 4;	Length 1463;
Best Local Similarity	98.5%	Pred. No. 0;		
Matches 1435; Conservative	0;	Mismatches 22;	Indels 0;	Gaps 0;

OY	19	AAIACCTGGCGGTGCGGCTTCTTCGCCCACCTCCGGCTACGCGCCGCCGCCACCCGCC	78
	1		
	2		
	3		
	4		
	5		
	6		
	7	AATACCTGCGCGCCCGCGGTTCTTCTTCGCGCACCTCCGGCTACGCGCCGCCGCCACCCGCC	66
OY	79	ACAGCGCGCGCAGACGAGATGGGCATCAAGGGTTTGAGAAACTCTGCGGGACAAATGCG	138
	1		
	2		
	3		
	4		
	5		
	6		
	7	ACAGCGCGCGCAGACGAGATGGGCATCAAGGGTTTGAGAAACTCTGCGGGACAAATGCG	126
OY	139	CCCAAGGGCGATGAAGAGACAGCAAGTTTCAGAGACTACTTCGCGCCGCAAAATGCGCGCTGCAC	198
	1		
	2		
	3		
	4		
	5		
	6		
	7	CCCAAGGGCGATGAAGAGACAGCAAGTTTCAGAGACTACTTCGCGCCGCAAAATGCGCGCTGCAC	186
OY	199	GCCAGCATGAGCATCTACCACTTCTCTGATAGTAGTTGCAAGACAGCGCATGSAAACTCTC	258
	1		
	2		
	3		
	4		
	5		
	6		
	7	GCCAGCATGAGCATCTACCACTTCTCTGATAGTAGTTGCAAGACAGCGCATGSAAACTCTC	246
OY	259	ACAAATGGAAGCGTGGAAGTCACTGATATTGCAAGGATTTTCACCGGCAATTAAGA	318
	1		
	2		
	3		
	4		
	5		
	6		
	7	ACAAATGGAAGCGTGGAAGTCACTGATATTGCAAGGATTTTCACCGGCAATTAAGA	306
OY	319	TTACTGGAAGCGGGGAATCAAGCCAGTTTATGTTTGTATGGCAAGCTCTCGATATGAAG	378
	1		
	2		
	3		
	4		
	5		
	6		
	7	TTACTGGAAGCGGGGAATCAAGCCAGTTTATGTTTGTATGGCAAGCTCTCGATATGAAG	366
OY	379	AAACAAGAACTTGCCTAAAGAATACCTCAAAAAGAGATGATGCACAACCAAGATGTGACTGAG	438
	1		
	2		
	3		
	4		
	5		
	6		
	7	AAACAAGAACTTGCCTAAAGAATACCTCAAAAAGAGATGATGCACAACCAAGATGTGACTGAG	426
OY	439	GCAGTAGAGAGTGAAGATTAAGATCGGATTGSAAAATTGAGCAACAGAGACTGTAAAGTGC	498
	1		
	2		
	3		
	4		
	5		
	6		
	7	GCAGTAGAGAGTGAAGATTAAGATCGGATTGSAAAATTGAGCAACAGAGACTGTAAAGTGC	486
OY	499	ACAAGGCAACACACACGAAGATTGTAACGACTATTAAAGCTTTAAGACTTATGGGGTTCTCTTTGTA	558
	1		
	2		
	3		
	4		
	5		
	6		
	7	ACAAGGCAACACACACGAAGATTGTAACGACTATTAAAGCTTTAAGACTTATGGGGTTCTCTTTGTA	546
OY	559	GAGGCACCTTCTGAAGCAGACCAAGATGTGAGCCCTTTCGATTAAGCATTAAGGTGTTCC	618
	1		
	2		
	3		
	4		
	5		
	6		
	7	GAGGCACCTTCTGAAGCAGACCAAGATGTGAGCCCTTTCGATTAAGCATTAAGGTGTTCC	606

QY	619	GCTGTGCTTAGAAGATTAAGAGACTCCCTACTTTTGGGGCTCCAGCTTCCTGCTAT	678
QY	607	GCTGTGCTTAGAAGATTAAGAGATTAAGAGACTCCCTACTTTTGGGGCTCCAGCTTCCTGCTAT	666
QY	679	TTAATGATCCAACTTCCAGAAATATACCTGTGATGGAAATTTGATGTTGCCAAGGTTTG	738
Db	667	TTAATGATCCAACTTCCAGAAATATACCTGTGATGGAAATTTGATGTTGCCAAGGTTTG	726
QY	739	GAGAGACTTGAACCTCACCATGACACAGCTCATTTGATTCATCCTGTGTGATGTAC	798
Db	727	GAGAGACTTGAACCTCACCATGACACAGCTCATTTGATTCATCCTGTGTGATGTAC	786
QY	729	TATTGTGTATACATCAAAAGGATATCGGGGGGCAACAGCTCTGAACCTTATTTGTCACAT	858
Db	787	TATTGTGTATACATCAAAAGGATATCGGGGGGCAACAGCTCTGAACCTTATTTGTCACAT	846
QY	859	GGGTCCATAGAAAGCATCTTGAGAAATCTTTAATAAAGACAGATATCAAAATTCCTGAGAC	918
Db	847	GGGTCCATAGAAAGCATCTTGAGAAATCTTTAATAAAGACAGATATCAAAATTCCTGAGAC	906
QY	919	TGCGCTTACCAAGAGCTCGACGCTTGTTCAAGAGCCTTAATGTCATTGATATTCCT	978
Db	907	TGCGCTTACCAAGAGAGCTCGACGCTTGTTCAAGAGCCTTAATGTCATTGATATTCCT	966
QY	979	GAGCTAAATATGATTCGACCTGATATGAGGAGGGCTCATTAAGTTTCCCTGTAATAAATAT	1038
Db	967	GAGCTAAATATGATTCGACCTGATATGAGGAGGGCTCATTAAGTTTCCCTGTAATAAATAT	1026
QY	1039	GCTTTCATGATGATCGGGGTGACAAAGGCCATAGAGATCAAAATCTGCCAAGATATA	1098
Db	1027	GCTTTCATGATGATCGGGGTGACAAAGGCCATAGAGATCAAAATCTGCCAAGATATA	1086
QY	1099	TCGTGCGAAGGAAAGACTCGAGTCTTTTAAAGCCACTGCAACATCAGCACCCTTA	1158
Db	1087	TCGTGCGAAGGAAAGACTCGAGTCTTTTAAAGCCACTGCAACATCAGCACCCTTA	1146
QY	1159	AAACGGAAGAGACTTGGGATTAATAACCAAGCAAGCAGCTGCGAACAAAGAAACAAAGCT	1218
Db	1147	AAACGGAAGAGACTTGGGATTAATAACCAAGCAAGCAGCTGCGAACAAAGAAACAAAGCT	1206
QY	1219	GGTGGAAGAAAGAAATATATCTTGGATGCTTGATGTACAACACTCAGACTACGAAGCAGGG	1278
Db	1207	GGTGGAAGAAAGAAATATATCTTGGATGCTTGATGTACAACACTCAGACTACGAAGCAGGG	1266
QY	1279	TGGCATGATACCTTGGCTTAGATTATTTAACTCCCTGTTTAACTCAGACTTTGTGTAA	1338
Db	1267	TGGCATGATACCTTGGCTTAGATTATTTAACTCCCTGTTTAACTCAGACTTTGTGTAA	1326
QY	1339	AGTTTGGCCATGTTCAAGCTGGGGGTAAGTAGTTGTGTTAAGAGANTGGGTGACAA	1398
Db	1327	AGTTTGGCTCATGTTTCAAGCTGGGGGTAAGTAGTTGTGTTAAGAGANTGGGTGACAA	1386
QY	1399	GTAACAACACTATTCGCTGTTTTTAACTCTGTCTCTTGAAGATATGATGCCACTATAA	1458
Db	1387	GTAACAACACTATTCGCTGTTTTTAACTCTGTCTCTTGAAGATATGATGCCACTATAA	1446
QY	1459	AAAAAAAAAAAAAAAAAAAA 1475	
Db	1447	AAAAAAAAAAAAAAAAAAAA 1463	
<p>           RESULT 3            US-09-426-557-3            : Sequence 3, Application US/09426557            : Patent No. 623527            : GENERAL INFORMATION:            : APPLICANT: Mahajan, Pramod B.            : TITLE OF INVENTION: Maize Rad2/FEN-1 orthologues and uses            : FILE REFERENCE: 0961            : CURRENT APPLICATION NUMBER: US/09/426,557            : CURRENT FILING DATE: 1999-10-22         </p>			

```

? EALIER APPLICATION NUMBER: 60/1112,332
? EALIER FILING DATE: 1998-12-15
? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: FastSeq for Windows Version 3.0.0
? SEQ ID NO 3
? LENGTH: 1541
? TYPE: DNA
? ORGANISM: Zea mays
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (79)...(1215)
US-09-426-557-3

```

Query Match	Score	DB 4	Length
94.88;	1401.2;		1541;

Matches 1409; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY	33	GGGTTTCCTWGGCCACGCGGCTCAGCGCGCGCCGCGCACCGCGCACGCGCGCGAGA	92
Db	15	GCGGTTTCTTGCCGACCTCCGGCTCAGCGCGCGCCGCGCACCGCGCACGCGCGCGAGA	74
OY	93	CGAGATGGGCGATCAAGCGGTTTGACGAAACTCTGGCGGACANTCGGCCAAGCGGATGAA	152
Db	75	CGAGATGGGCGATCAAGGGGTTTGACGAAACTCTGGCGGACANTCGGCCAAGCGGATGAA	134
OY	153	GGAGCAGAAGTTTCGAGAGCTACTTCGGCCGAAANTCGCGCTGCAGCGCGACATGACAT	212
Db	135	GGAGCAGAAGTTTCGAGAGCTACTTCGGCCGAAANTCGCGCTGCAGCGCGACATGACAT	194
OY	213	CTACCACTTCCGATAGTAGTTGGAAAGCAGCGCATGGAAACTCTGCACAAATGAAGCTGG	272
Db	195	ATACCACTTCCGATAGTTAGTTGGAAAGCAGCGCATGGAAACTCTGCACAAATGAAGCTGG	256
OY	273	TGAAGTCTACTAGTCATTTGACAGGAATGTTCAACCGCAATAAGATTACTGGAAGCGG	332
Db	255	TGAAGTCTACTAGTCATTTGACAGGAATGTTCAACCGCAATAAGATTACTGGAAGCGG	314
OY	333	AATCAGGCGAGTTATGTGTTTGGANTGGCAAGCCTTCGTGATTGAAAGAAACAAGACTTC	392
Db	315	AATCAGGCGAGTTATGTGTTTGGANTGGCAAGCCTTCGTGATTGAAAGAAACAAGACTTC	374
OY	393	TAAAGATACTCAAAAAGAGATGATGCACCAAGAATCTGACTAGGACAGTAGAGTAGG	452
Db	375	TAAAGATACTCAAAAAGAGATGATGCACCAAGAATCTGACTAGGACAGTAGAGTAGG	434
OY	433	AGATTAAGATCGGATGTAATAATTGAGCAGAGAGACTGTAAAGTGCACAGGCAACAA	512
Db	435	AGATTAAGATCGGATGTAATAATTGAGCAGAGAGACTGTAAAGTGCACAGGCAACAA	494
OY	513	CGAATCTGTAAAGGACATTTAAGCTTAAGGGGGTTCCTGTGTAAAGGACCTTCTGA	572
Db	495	CGAATCTGTAAAGGACATTTAAGCTTAAGGGGGTTCCTGTGTAAAGGACCTTCTGA	554
OY	573	AGCACAAGCAGATGTGCGAGCCCTTTGGCATAAAGATTAAGTGTTCGCTTTCCTCAGA	632
Db	555	AGCACAAGCAGATGTGCGAGCCCTTTGGCATAAAGATTAAGTGTTCGCTTTCCTCAGA	614
OY	633	AGATTAAGGATCTCCCTACTTTTGGGGGCTCACGGTTCCCTTGTCATTTAATGATCCAG	692
Db	615	AGATTAAGGATCTCCCTACTTTTGGGGGCTCACGGTTCCCTTGTCATTTAATGATCCAG	674
OY	693	TTCCAAGAAATTAACCTGTGATGAGTAATTTGATGTCGCAAGGTTTGGAGGAGCTTGACT	752
Db	675	TTCCAAGAAATTAACCTGTGATGAGTAATTTGATGTCGCAAGGTTTGGAGGAGCTTGACT	734
OY	753	CACCATGAGCAGTTCATTTGATTTGTGCATCTGTGTGATGTGACATTTGTGATAGCAT	812
Db	735	CACCATGAGCAGTTCATTTGATTTGTGCATCTGTGTGATGTGACATTTGTGATAGCAT	794
OY	813	CAAAAGCTTGGGGGGGCAACAAGCTGTGAACCTTATTCGTACACATGGGTCATAGAAAG	872
Db	795	CAAAAGCTTGGGGGGGCAACAAGCTGTGAACCTTATTCGTACACATGGGTCATAGAAAG	854

OY	873	CATCTTGGAGATCATCTTAATAAAGACAGATATCAAAATTCCTGAGGACTGCCTTTCCAAAGA	932
Db	855	 CATCTTGGAGATCATCTTAATAAAGACAGATATCAAAATTCCTGAGGACTGCCTTTCCAAAGA	914
OY	933	AGCTCGACGCTTGTTTCAAAGAGCCTTAATGCATCTTTGGATATTCCGTGACCTAAAATGGAC	992
Db	915	AGCTCGACGCTTGTTTCAAAGAGCCTTAATGCATCTTTGGATATTCCGTGACCTAAAATGGAC	974
OY	993	TGCACCTCATATGAGAGGGGTCTCATPAGCTTTTCTCGTTAAAGATATATGTTTCATGAAGA	1052
Db	975	TGCACCTCATATGAGAGGGGTCTCATPAGCTTTTCTCGTTAAAGATATATGTTTCATGAAGA	1034
OY	1053	TCGGGTGACAAAGGCCATAGAGAGATCAATATCTGCCAAGATAAATGCTGCAAGAGAG	1112
Db	1035	TCGGGTGACAAAGGCCATAGAGAGATCAATATCTGCCAAGATAAATGCTGCAAGAGAG	1094
OY	1113	ACTGCAATCTTTTTTCAAAGCCAATCTGCCACAATCAGCACCGGCTAAAAAGSAGAGAC	1172
Db	1095	ACTGCAATCTTTTTTCAAAGCCAATCTGCCACAATCAGCACCGGCTAAAAAGSAGAGAC	1154
OY	1173	TTTCGATTTAAACCAACCAAGGACGCTGGGACACAAAGAAAGGTTGGTGGGAAGAAAGA	1232
Db	1155	TTTCGATTTAAACCAACCAAGGACGCTGGGACACAAAGAAAGGTTGGTGGGAAGAAAGA	1214
OY	1233	ATAATCTTGGATGCTTGTATGTATGATACAACATCAGACTACGAAAGACGCGTGGCATGTACTT	1292
Db	1215	ATAATCTTGGATGCTTGTATGTATGATACAACATCAGACTACGAAAGACGCGTGGCATGTACTT	1274
OY	1293	GCGCTAGATATTATTACCTCCCTGTTTAACTCAGACCCTTTGGTGAAGTTGGCCATGTT	1352
Db	1275	GCGCTAGATATTATTAAACCTCCCTGTTTAACTCAGAGCCTTTGGTGAAGTTGGCTATGTT	1334
OY	1353	TCAACCTGGGGTAAAGTATGTTGTGTTGAAGAGATTGGTGTACCAAGTAAACAAACTAT	1412
Db	1335	TCAACCTGGGGTAAAGTATGTTGTGTTGAAGAGATTGGTGTACCAAGTAAACAAACTAT	1394
OY	1413	CGCTGTTTTTTACTCTCTTGCTCCTTTGAAGATATGATVSCCACT	1454
Db	1395	CGCTGTTTTTTACTCTCTTGCTCCTTTGAAGATATGATVSCCACT	1436
RESULT 4			
US-09-426-557-5			
; Sequence 5, Application US/09426557			
; Patent No. 6232527			
; GENERAL INFORMATION:			
; APPLICANT: Mahajan, Pramod B.			
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses			
; FILE REFERENCE: Thereof			
; CURRENT APPLICATION NUMBER: US/09/426.557			
; CURRENT FILING DATE: 1999-10-22			
; EARLIER APPLICATION NUMBER: 60/112.332			
; EARLIER FILING DATE: 1998-12-15			
; NUMBER OF SEQ ID NOS: 10			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 5			
; LENGTH: 1381			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (37)...(1173)			
; US-09-426-557-5			

[illegible]

QY	121	CTGTGGGGGACATGGCCCAAGGGGATGAAGSACACAGCTTGAGAGCTACTTGGC	180
Db	61	CTGTGGGGGACATGGCCCAAGGGGATGAAGSACACAGCTACTTGGC	120
QY	181	CGCAAAATCGCGCTGCAAGCCAGCATGAGATCTACAGTTCCTGATGATGGAAG	240
Db	121	CGCAAAATCGCGCTGCAAGCCAGCATGAGATCTACAGTTCCTGATGATGGAAG	180
QY	241	ACAGGACGTGAACCTCTCAAAATGAAAGCGGTGAAGTCATGATCATTTGCAAGATG	300
Db	181	ACAGGACGTGAACCTCTCAAAATGAAAGCGGTGAAGTCATGATCATTTGCAAGATG	240
QY	301	TTTCACCGGACAAATAGATTAAGTACTGGAAGCGGATCAAGCCAGTTATGTTTTGATGGC	360
Db	241	TTTCACCGGACAAATAGATTAAGTACTGGAAGCGGATCAAGCCAGTTATGTTTTGATGGC	300
QY	361	AAGCCTCTCTGATATGAAGAAACAAGAAGCTTGCATAAAAGTACTCAAAAAGATCATGCA	420
Db	301	AAGCCTCTCTGATATGAAGAAACAAGAAGCTTGCATAAAAGTACTCAAAAAGATCATGCA	360
QY	421	ACCAAGATCTGACTGAGCGAGTGAAGTGAAGATGAAGATGCGATGCAAAAATTGAGC	480
Db	361	ACCAAGATCTGACTGAGCGAGTGAAGTGAAGATGAAGATGCGATGCAAAAATTGAGC	420
QY	481	AAGGAGCTGTAAGGTCACAGGCAACACACGAAGATTGTAAAGACTATTAAGCTT	540
Db	421	AAGGAGCTGTAAGGTCACAGGCAACACACGAAGATTGTAAAGACTATTAAGCTT	480
QY	541	ATGGGGTTCCTCTTTGTAAGGCAACCTTCGAAGCAAGACGAATGTCAACCTTTGCG	600
Db	481	ATGGGGTTCCTCTTTGTAAGGCAACCTTCGAAGCAAGACGAATGTCAACCTTTGCG	540
QY	601	ATAAACGATAAGGTTGCGCTGTGCTTCGAAGATTAAGGACTCCCTTACTTTGGGCT	660
Db	541	ATAAACGATAAGGTTGCGCTGTGCTTCGAAGATTAAGGACTCCCTTACTTTGGGCT	600
QY	661	CCAGGTTCCCTGCTCATTTAATGATCCAGTTCACAGAAATACCTGATGTAATTT	720
Db	601	CCAGGTTCCCTGCTCATTTAATGATCCAGTTCACAGAAATACCTGATGTAATTT	660
QY	721	GATGTTGGCAAGTTTTGGAGAGACTTGAACCTCACCATGGACCGATTTGATTTGTGC	780
Db	661	GATGTTGGCAAGTTTTGGAGAGACTTGAACCTCACCATGGACCGATTTGATTTGTGC	720
QY	781	ATCCTGTGTGATGTGACTATTGTGATAGATCAAAAGTATCGGGGGCAACACTCTG	840
Db	721	ATCCTGTGTGATGTGACTATTGTGATAGATCAAAAGTATCGGGGGCAACACTCTG	780
QY	841	AAACTTATTCCTCAACATGGGCTCATGGAAGCATTTGGAGATTTTAAATAAGACGA	900
Db	781	AAACTTATTCCTCAACATGGGCTCATGGAAGCATTTGGAGATTTTAAATAAGACGA	840
QY	901	TATCAATTCCTGAGACTGGGCTTACCAAGAAGCTGAGGCTTTGTCAAGAGCTTAT	960
Db	841	TATCAATTCCTGAGACTGGGCTTACCAAGAAGCTGAGGCTTTGTCAAGAGCTTAT	900
QY	961	GTCACATTGGATTTCCGTGAGACTTAAATGAGCTGCACTGATAGAGAGGCTTCATAACT	1020
Db	901	GTCACATTGGATTTCCGTGAGACTTAAATGAGCTGCACTGATAGAGAGGCTTCATAACT	960
QY	1021	TTTCTGTGTAAGATTAATGTGTTTCATGAAGATCGGGTGCAAAAGGCCATAGAGAATC	1080
Db	961	TTTCTGTGTAAGATTAATGTGTTTCATGAAGATCGGGTGCAAAAGGCCATAGAGAATC	1020
QY	1081	AAATTCGCCAAGATTAATGCTGCGCAAGAGACTCGAGTCTTTTTCAGGCAACTGGC	1140
Db	1021	AAATTCGCCAAGATTAATGCTGCGCAAGAGACTCGAGTCTTTTTCAGGCAACTGGC	1080
QY	1141	ACCACATCAGACCGCTTAAAGGGAAGACATTCGATTAACCAAGCAAGGCACTGTGG	1200
Db	1081	ACCACATCAGACCGCTTAAAGGGAAGACATTCGATTAACCAAGCAAGGCACTGTGG	1140
QY	1201	AACAAGAAACAAGGCTGTGTGAAGAAAGAAATATCTTGATGCTGTGATGACACTTA	1260



|||||  
Db 1141 AACAGAAACAAAGGCTGGTGAAGAAAGAAATCTGGATGCTGATGACAACTA 1200  
QY 1261 CGACTAGAAAGACGCGGTGATGATCATCTGCGCTAGATTATTTAACTCCCTGTTTA 1320  
Db 1201 CGACTAGAAAGACGCGGTGATGATCATCTGCGCTAGATTATTTAACTCCCTGTTTA 1260  
QY 1321 ACTGAGACCTTTGGTGAAGTTTGCCCATGTTTCAGCTGGGTAGTAACTGTTGTTG 1380  
Db 1261 ACTGAGACCTTTGGTGAAGTTTGCCCATGTTTCAGCTGGGTAGTAACTGTTGTTG 1320  
QY 1381 AAGGATTGGTGTACCAAGTACAAACTATCGCTGTTT 1423  
Db 1321 AAGGATTGGTGTACCAAGTACAAACTATCGCTGTTT 1363

RESULT 5  
US-08-455-968E-2  
; Sequence 2, Application US/08455968E  
; Patent No. 5874283  
; GENERAL INFORMATION:  
; APPLICANT: Harrington, John L.  
; APPLICANT: Hsieh, Chih-Lin  
; APPLICANT: Lieber, Michael  
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/455,968E  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 18985-000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1144 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-08-455-968E-2

Query Match 20.3%; Score 299.4; DB 2: Length 1144;  
Best local Similarity 56.5%; Pred. No.4.6e-73;  
Matches 599; Conservative 0; Mismatches 456; Indels 6; Gaps 2;

QY 97 ATGGGATCAAGGTTTGAGCAATGCTGCGGACAAATCGCGCAAGCGGATGAAGAG 156  
Db 1 ATGGGATCAAGGTTTGAGCAATGCTGCGGACAAATGCTGCGGACAAATCGCGGAG 60  
QY 157 CACAAGTTCAGAGCTACTTGGCCGCAAAATGCGCGTCGACGCCACATGAGCATCTAC 216  
Db 61 AATGACATCAAGAGCTACTTGGCCGCAAAATGCGCGTCGACGCCACATGAGCATCTAT 120  
QY 217 CATTGCTGATAGTGTGAGAGACAGCAAGTCAAACTCTCAAAATGAAGTGGTGA 276  
Db 121 CATTGCTGATAGTGTGAGAGACAGCAAGTCAAACTCTCAAAATGAAGTGGTGA 177

QY 277 GTCACTAGTCATTTCGAAGAATGTTCAACCGGACAAATAGATTACTGGAAGCGGAATC 336  
Db 178 ACCACACACCCACGATGGGATGTTCTACCGCACATTCGATGATGAACAGGCAATC 237  
QY 337 AAGCCAGTTATGTTTTTGTATGGAAGCCCTCTGATATGAAGAAACAAAGAACTTCTAAA 396  
Db 238 AAGCCAGTTATGTTTTTGTATGGAAGCCCTCTGATATGAAGAAACAAAGAACTTCTAAA 297  
QY 397 AGATATCTCAAAAGATGATGACCAACCAAGATCTGACTGAGCAATGAGTGAAGAT 456  
Db 298 CGCAGTGAAGCGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 357  
QY 457 AAGATGCGATTGAAAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516  
Db 358 GAGCAGAGAGTGGAAAATTTACTAAGCGGCTGTTGAAGTCTACTAGACAGCAATGAT 417  
QY 517 GATTGTAAAGCACTATTAAGACTTATGAGGAGTCTCTGTTTGAAGGACCTTCTGAAGCA 576  
Db 418 GAGTGAAGCACTATGAG 477  
QY 577 GAGCAGAGATGTGAGCCCTTTCATTAAGATTAAGTGTGCTGTTGCTTCAAGAGAT 636  
Db 478 GAGCAGAGATGTGAGCCCTTTCATTAAGATTAAGTGTGCTGTTGCTTCAAGAGAT 537  
QY 637 AAGGACCTCCCTTACTTTTGGGGGCTCAAGGTTCTGCTGCTTAAATGATCCAAAGTTCC 696  
Db 538 ATGACCTGCTCCTACCTTCGAGAGCCCTGCTGCTTAAATGACACCTGACTGCAAGAGCC 597  
QY 697 AAGAAATACCTGTGATGAATTTGATGTTGCAAGGTTTGGAGAGAGCTTGAAGTCAAC 756  
Db 598 AAAAAGCTGCAATCCAGAGATTCACCTGAGCCGAGATTCGAGAGAGCTGAGCTGAG 657  
QY 757 ATGACACATTCATGATTTTGTGATCTGCTGTGATGATGATGATGATGATGATGAT 816  
Db 658 CAGGAACAGTTTGTGATCTGCTGTGATGATGATGATGATGATGATGATGATGATG 717  
QY 817 GGTATCGGGGGGCAAAACACTTAACTTATGCTCAACCTGAGTCAATGAAAGATC 876  
Db 718 GGTATCGGGGGGCAAAACACTTAACTTATGCTCAACCTGAGTCAATGAAAGATC 777  
QY 877 TTGAGAGATCTTAATTAAGACAGATATCAAAATTCCTGAGAGCTGCTTACCAAGAGCT 936  
Db 778 GTGCGGAGACTTGAACCCCAACAGTACCTGCTGAGAGAAATTCGCTCCCAAGAGAGCT 837  
QY 937 CGAGCTTGTTCAGAGAGAGCTAATG---TACTTTGATATTCCTGAGCTAATGAGACT 993  
Db 838 CACGAGCTTGTTCAGAGAGAGCTAATG---TACTTTGATATTCCTGAGCTAATGAGACT 897  
QY 994 GCACCTGATGAGAGAGGCTCATAGTTTCTGTTAAAGTAAATGTTCAATGAAGAT 1053  
Db 898 GAGCAATTAAGAGAGAGCTGATCAAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTG 957  
QY 1054 CGGTTGACAAAGGCCATAGAGAGATCAAAATCTGCCAAGATTAATCTGTCGAGAGAGA 1113  
Db 958 CGAATTCGAGAGTGGGTTGAGAGAGCTGAGTAAAGAGCCGCCAAGAGCAGACCCAGAGCG 1017  
QY 1114 CTGAGATCTTTTTCAGGACCACTGCCACATCAGCAGC 1154  
Db 1018 CTGAGATCTTTTTCAGGACCACTGCCACATCAGCAGC 1058

RESULT 6  
US-08-455-968E-28  
; Sequence 28, Application US/08455968E  
; Patent No. 5874283  
; GENERAL INFORMATION:  
; APPLICANT: Harrington, John L.  
; APPLICANT: Hsieh, Chih-Lin  
; APPLICANT: Lieber, Michael  
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESSES:



```

:
:   TOPOLOGY: linear
:   MOLECULE TYPE: DNA (genomic)
:   FEATURE:
:   NAME/KEY: CDS
:   LOCATION: 104..1237
:   US-08-455-968E-9
:
Query Match      18.6%; Score 275; DB 2; Length 2033;
Best Local Similarity 55.8%; Pred. No. 3,3e-66;
Matches 637; Conservative 0; Mismatches 480; Indels 24; Gaps 5;

QY  97  ATGGGCAATCAAGGGTTTGACGAACACTGCTGGCGGACATATGGGCCCAAGCCATGAGAG 156
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db  104  ATGGAATTCACGGCCTTGCCAAACTAATTCGTATGTGGCCCGCCAGTCATCCGTAG 163
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY  157  CAGAAGTTGAGAGCTTACTTGGCCGCAAAATCGCCGTGAGCGCAGCATGAGATCTAC 216
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db  164  AATGACATCAAGACTACTTTGGTCGCAAAAGTGGCCATCGATGCGCTCATGACATCTAC 223
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY  217  CAGTTCTGATAGTAGTTGGAGAGCAGGCGATGCAACTCTCACAAATGAAGCTGGTGA 276
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  224  CAGTTCTGATAGTAGTTGTCGACAG--GGTGGGATGTGTGACAGAAAGAGAGGCTGAG 280
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  277  GTCACTAGTCAATTGGCAAGATGTTCAACGGGACAAATAGATTACTGAGAGCGGGATC 336
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  281  ACCACACAGCC---TGATGGGCATTTTCAACGTCATCGC---CATGGAAATGGCATC 334
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  337  AAGCCAGTTATGTTTTTGTATGAGCAGCCTCTGATATGAAGAAACAGAACTTGCTAA 396
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  335  AAGCCTGTGCTGCTTTGATGAGCAAAACACACAGCTGAAGTGAAGGAGACTGGCCAG 394
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  397  AGATACATCAAAAAGATGATGACCAACCAAGATCTGACTGAGGAGTAGAGCTAGGAGAT 456
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  395  CGCAGTAGAGGCGCGCGGAGGCTGAGAGAGCAAGCTGACAGCGCTCAGCAGCTGGGATG 454
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  457  AAAGATCGATTGAAATATTGACAGAGAGCATGTAAAGGTCACAGAGCAACACAGAA 516
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  455  GAGAGAGAGGTGAGAGATTCCACAGAGAGGCTGTGAAGGTACACCAAGCAACACAAATGAT 514
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  517  GATTGTAAACGACTATTAAAGACTTATGGGGTCTCTGTTAGAGGACCTTTCAGCA 576
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  515  GAGTGCAAAACACCTGCTGAGCGCTCATGGGACATCCCTTACCTTGTATGACCCAGGAGGCA 574
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  577  GAAGCAAAATGTGAGGCGCTTTCATTAAGCATTAAGGTGTTGCTGTGCTTCAAGAT 636
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  575  GAGGCCAGCTGTGCTGCGCCGCAAGAGCTGGCGCAAGCTATGCTGGCGCCAGGAGGAC 634
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  637  AAGGACTCCCTTACTTTTGGGGGCTCCACGGTTCCTGTCATTTAATGATCCAAGTTCC 696
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  635  ATGGACTGCTCACTTTTGGCAGGCCCGTGTATGCGACACTTAACCTGCCAGTGAAGCC 694
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  697  AAGAAATACCTGTGATGAATTTGATGTTGCCAAGGTTTTTGAAGAGCTTGAAGTCAAC 756
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  695  AAGAAAGCTGCCCATCCAAAGATTCCATCTAGAGCGCTCCCTGCGAGAGCTGGGCTGAAC 754
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  757  ATGGACCAATTCAATTTTGTGATCTCTGTGATGATGATGATGATGATGATGATGATGAT 816
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  755  CAGGAGAGATTTTGTGATCTGTGATCTCTGTGATGATGATGATGATGATGATGATGAT 814
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  817  GGTATCGGGGGGCAAAACAGCTCTGAACCTTATTCGTCACATGATGGGTCCATAGAAAGCAT 876
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  815  GGCATTGGCCCAAGCGGGGCTGTGGATCTCATCCAGAAACATAAGACATCGAGAGATC 874
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  877  TTGGAGATCTTATTAAGACAGATATCAAAATTCCTGAGAGCTGGCCTTACCAAGAAGCT 936
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  875  GTAGAGGGGTGAGACCCACCAAGTACCCGTTCCAGAGAACTGGCTCCACAAAGAGGCC 934
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  937  CGAGGCTTTTCAAGAGGCTAA---TGTCACCTTTGATATTTCTCTGAGCTTAAATGCACT 993
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  935  CACAGACTCTTCTCGAGAGCAGAAATGATGATGATGATGATGATGATGATGATGATGAT 994
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  994  GCACTGTAGTAGAGAGGCTCATTAAGTTTCTCTGTAAAGATATGATGATGATGATGAT 1053
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

```

Db  995  GAGCCAAATGAGAGAGAGTTGCTCAATTTATGTGTGTAAGAAAGCAGTTTTTGAAGAG 1054
QY  1054  CGGGTGACAAAGGCCATAGAGATGATCAATCTGCCAGAAATTAATTCGTCCGACAGAGA 1113
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1055  CGAATTTCCAGTGGGGGTCAAGGGGCTGAGTAAAGAGCCCGCAGGAGCAGCACCAGGGAGCG 1114
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  1114  CTCGAGTCCCTTTTTCAGCCCACTGCCCACACATCAGCAGCGCTAAAGACGAAGGAGACT 1173
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1115  CTCGATGATTTCTTCAAGGTGACAGGCTCAGCTCTCTCAGC-----TAAAGCGC 1162
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  1174  TCGGATTAACACAGACAGCAGCTCGACACAGAAACAAAGCGCTGTGGAAAGAGAAA 1233
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1163  AAGGAGCCAGAACCCAGGGGCTGCTAAGAAAGAACAAAGACTGGGGAGAGCGGGAAG 1222
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY  1234  T 1234
Db  1223  T 1223
```

```

RESULT 8
US-08-455-968E-4
: Sequence 4, Application US/08455968E
: Patent No. 5874283
:
: GENERAL INFORMATION:
:   APPLICANT: Harrington, John L.
:   APPLICANT: Hsieh, Chih-Lin
:   APPLICANT: Lieber, Michael
:   TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
:   NUMBER OF SEQUENCES: 63
:
: CORRESPONDENCE ADDRESSES:
:   ADDRESSEE: Townsend and Crew LLP
:   STREET: Two Embarcadero Center, 8th Floor
:   CITY: San Francisco
:   STATE: California
:   COUNTRY: USA
:   ZIP: 94111-3834
:
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/455,968E
:   FILING DATE: 30-MAY-1995
:   CLASSIFICATION: 435
:   ATTORNEY/AGENT INFORMATION:
:     NAME: Apple, Randolph T.
:     REGISTRATION NUMBER: 36,429
:     REFERENCE/DOCKET NUMBER: 18985-000100
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: 415-576-0200
:       TELEFAX: 415-576-0300
:       INFORMATION FOR SRO ID NO: 4:
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 1930 base pairs
:       TYPE: nucleic acid
:       STRANDEDNESS: single
:       TOPOLOGY: linear
:     MOLECULE TYPE: cDNA
:
US-08-455-968E-4

Query Match      18.3%; Score 270.6; DB 2; Length 1930;
Best Local Similarity 55.0%; Pred. No. 5.2e-65;
Matches 628; Conservative 0; Mismatches 489; Indels 24; Gaps 4;
```

OY	217	CATTCCGTGATGATGTTGGAAAGGACAGCGCATGGCAAACTCCACCAATGAAAGCTGGGGAA	27
Db	121	CAGTTCCTGATGCTGTTCGTCAAG--GGTGGGATGTGCTGCAGAACGACGAGAGGTGAG	177
OY	277	GTCACCTAGTCATTTGGCAAGGAATGTCACCGGACCAATAAGATTACTGGAAAGCGGGATC	336
Db	178	ACCAACAGCCCTGATGGGATGTTAT-----GGCAAAACCATCCGCATGAGGAATGGCATC	231
OY	337	AAGCCAGTTTATGTTTTTGATGGCAAGCCTCTGTATATGAAGAAACAAGAACTTGTCTAA	356
Db	232	AAGCTGTGTACGTCCTTGATGGGCAAAACACACACAGCTGAAGTCAAGCGAGCTGGCCAA	291
OY	397	AGATACCTCAAAAAGATGATGTCAAACCAAGATGTGACTGAGGCAATGATGATAGCAGAT	456
Db	292	CGCAGTGAAGAGCCGCCGAGAGCTGAGAGAGCAACTGCAGCAGGCTCAGAGAGCTGGGATG	351
OY	457	AAAGATGCGATTGAAAAATTGAGCAAGAGCAGTGTAAAGTCCACAGGCAACACACGAA	516
Db	352	GAGGAGGAGTGGGAGAAATTCCACCAAGAGCTCGTGAAGGTCACACAGCAACCAATGAT	411
OY	517	GATGTAAAGCAATTAAGACTATGGGGGCTCTGTGTGAAGGAGCACTTTCGAAGCA	576
Db	412	GAGTCAAAACACCTGTGTAGGCTCATGGGCATCCCTTACTGTGATGCAACCCAGCGAGCA	477
OY	577	GAAGCAGATGTGCAAGCCCTTTGGCATAAACATTAAGTGTGTGCTGTGCTTCGAAGAT	636
Db	472	GAGGCGACGCTGTGGTGGCCCTGGCAAAAGCTGGCAAAAGTCTATCTGCGCGCCACGAGAG	531
OY	637	AAGGACTCCCTTACTTTTGGGGGCTCCAGAGTTCTCTGTGATTAATGATCCAAATTC	696
Db	532	ATGGACTCCTCCTACTTTTGGAGAGCCCGCTGTAAATGGACCTTAAGTCCAGTAGAGCC	591
OY	697	AAGAAATACCTGTGATGCAATTTGATGTTGCCAAGGTTTGGAGAGAGCTGGAATCACC	756
Db	592	AAGAAGCTGCCCATCCAGAGATTCATCTGAGCCGGCTCCGACAGAGCTGGGTCTAAC	651
OY	757	ATGGAACCACTTATGATTTGTGCATCTGTGTGATGTACTATTGTGATAGCATCAA	816
Db	652	CAGGAGCAGTTGTGTGATCTGTGATCCTGTGTGGGTACCGACTCTCGAGACGATCCGT	711
OY	817	GGTATCGGGGGGCAACAGCTCTGAAACTTATTTGTCMAACTGTGATCAGTGAAGCAATC	876
Db	712	GGCATTTGGCGCCMAAGCGGCTGTGATCTTCATCCAGAAACATTAAGACATGAGAGATC	771
OY	877	TTGGAAGATCTTAAATTAAGACAGATATCAAAATTCCTGAGAGCTGGCTTACCAAGAACT	936
Db	772	GTAAGAGCGGCTGAGACCCAGCAAGTAAACCCCTGTCCAGAGAACTGGCTCCACAAAGAAC	831
OY	937	CGAGCCTGTGTCAGAGAGCCCTAA--TGTCACCTTGGATATTCCTGAGCTTAAATGGACT	992
Db	832	CAGCAGCTCTTCTGTGAGACCGAGAACTAGTGACACCCAGAGTGTGTGAGACTTAAATGAGC	891
OY	994	GCACCTGATGAGAGAGGCTCATTAAGTTTCTGTGTAAGAAAGTAATGTTTCAATGAAGAT	1051
Db	892	GAGCCAAATGAAGAAAGAGTGTGTAATTTATGTGTGTAAGAAAGCAAGTTTCTGAAAGAG	951
OY	1054	CGGGTGACAAGAGCCATGAGAAAGTCAAACTGTGCCAAGAAATAATGCTGCCAAGGAGA	1111
Db	952	CGAATTCGACGTGGGGTCAAGCGGCTGTGATGAAGCCGCCAGGGCAAGCACCAGGAGCGC	1011
OY	1114	CTCGAGTCCCTTTTAAAGCAAACTGCCACCAATCAGCAGCCGTAAAGGAGAGCAACT	1171
Db	1012	CTCGATGATTTCTTAAAGTGAAGGCTCACTCTCTCAGC-----TAAGCGC	1071
OY	1174	TCGGATAAACAAGCAAGGCAAGCTGCGAACAAGAAACAAGGCTGTGGAAGAAGAAA	1231
Db	1060	AAGGAGCCAGAAACCCAAAGGGGCTGTAAAGAAAGAAAGAGCTGTGGGGAGAGCGGGAA	1131
OY	1234	T 1234	
Db	1120	T 1120	

```

RESULT 9
US-08-455-968E-6
; Sequence 6, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chin-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-May-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-455-968E-6

Query Match 18.0%; Score 265.4; DB 2; Length 1149;
Best Local Similarity 55.2%; Pred. No. 1.1e-63;
Matches 596; Conservative 0; Mismatches 456; Indels 27; Gaps 3;

QY 97 ATGGGCATCAAGGGTTTGACGAACCTGCTGGCGGACAAATGCGCCGCAAGCGCATGAAGAG 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 ATGGGATATTAAGGTTTGTGAATCAATTAATATGATGCAAGAACTGCTCCCTGCTATCAGAAA 60

QY 157 CAGAACTTGCAGAGCTACTTCGGCCCGCAAAATCCGCCCTGCAGCCGACGATGCGATCTAC 216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 AGCGATATCAAGAGCTTTTGGCAGAAAGGTTCGATCATGATGCTCTATATGTT 120

QY 217 CAGTTCCGTATAGTATGTTGGAGAGCAGCATGAGAACTCTCACAAAATGAAGCTGTGAA 276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 CAGTTTATATGCTGTGAAGACAGACGGGTGGGCAAGTTGACCAATGAAGCCGGTGA 180

QY 277 GTCACTAGTCATTGGCAAGAAATGTTCAACCGGACAAATGAATTAAGATTCTGGAAGCGGAATC 336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 ACAAGATCACTTATGATGGATGTTATGTTTATAGCACTGACATGAATGATTAACGGTATTC 240

QY 337 AAGCCAGTTATGTTTATGTTGATGGCAAGAGCTCTGATGAATGAAGAAACAGAACTTCTAAA 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 AAGCCCTGTTATGCTTCGACGGCABACCTCCAGCTTTGAATTCATGATGTCAGAAAG 300

QY 397 AGATACTCAAAAAGAGATGATGCAACCAAGAATCTGCAGTGCAGGCAAGTAGAGTAGAGAT 456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 CGGCTTCAAGAGGGTGGAAACAGAAAAAATACTGGCAGAGGCAACAACAGAAATTGAA 360

QY 457 AAGATGCAATTGAAAAAATTGACGAAGAGAGCACTGTAAGAGTACAAAGCAACACAGAA 516
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 AAGATGCAACCAAGAAAGAAAA-----ACATTGTTGAAGAGCTTCAAAAAGAGCTAATGAA 411

```

Tue May 27 16:02:32 2003

us-09-805-311-7.std.rni

Page 9

Qy	517	GATTTGAACAAGCATTTAAACATTTAAGGGGGTTCGTGTTGAGAGGACCTTCGTAAGA	576
Db	412	GAAGGCCAAAATTACTAGACACTAATGGGAATCCCATATATAAATAGCCCAACGAAAGCT	471
Qy	577	GAAGCAGAAATGCGACGCCCTTTGCAATAACGATAAGGTGTTGCGTGTCTCAGAAAGT	636
Db	472	GAGGCTCATGCTGTGAGTTGGCAAGAAAGGGAAGGTGTATGCGCGAGCAAGTGAAGAT	531
Qy	637	AAGGACTCCCTTACTTTTGGGGCTCCACGCGTCTCTCCATTAATAGATCCAAATGCC	696
Db	532	ATGGAGACACTCTGTATTAGAACACCCCTCTTGTTGAGACATTTGACTTTTCAGAGGCC	591
Qy	697	AAGAAATATCCGTGATGAGTAATTTGATTTGGCAGGTTTTGGAGAGCTTGAACTCAGC	756
Db	592	AAGAAGAAACCGATTACAGAAATAGATACGAATTAAGTTTGAAGAGACGTCGACTTGACA	651
Qy	757	ATGAGACAGTTCAATTGATTTGTGCATCTCGTGTGATGTGACTATTGTGATAGACAA	816
Db	652	ATGAGACAGTTGTTGATCTTTGCATATAAGCTGTGTTGTGACTCTGTGAACATCAACA	711
Qy	817	GGATGAGGGGGGCAAAACGCGCTGGAACATTAATCTGCACAGTGGGTGCATAGAAACACT	876
Db	712	GGTGTGGTCCAGTGACAGCCTTAATAATTGATTAATAAACGCTATGCATCAGAAAAAATC	771
Qy	877	TTGAGACATCTT-----AATAAGACAGATATCAAAATTCCTGAGAGCTGG	921
Db	772	GTGGAGTTTATATGATCTGGGGAGTCAAAACACACTTAATGGAATAATCCGAGAGACTG	831
Qy	922	CCTTACCAAGAAGCGCGAGCGTTGTTCCAGAGACCGTAATGTCACTTTGGATTA--TTGCT	978
Db	832	CCCTACCAAAACGAAGAAATGCTGTCTTTCTTGAGCCTGAAGTTATAGATGGTAAAGAAAT	891
Qy	979	GACCTAAATGACTGCACCTGATGAGAGGGTCTCATTAAGTTTCTGGTAAAGATTAAT	1033
Db	892	AACTTGAATGCTGCCACCAAGAGGAAGGAATTATCGAGTATTTATGTGATGATGAAG	951
Qy	1039	GGTTTCAATGAAGATCGGGGTGACAAAGCCATAGAGAATCATCAATCTGCAGAAATTA	1098
Db	952	AAATTCAGTGAAGAAAGATTAATCTGATATATCAATGATGAATAAAAGGCTTGAATCT	1011
Qy	1099	TGCTCCCAAGGAAGCTGAGATCCTTTTTCAGAACCACTGCCACACATCAGACCCGT	1157
Db	1012	GCGATTCAAGGTAGGTATGATGGTCTCTTCAGATGTTGGTGCCTAAGACAAAGAACAGCT	1070
US-RSULT 10			
US-08-757-653-175			
: Sequence 175, Application US/08757653			
: Patent No. 5843669			
GENERAL INFORMATION:			
APPLICANT: Kaiser, Michael W.			
APPLICANT: Lyamichev, Victor I.			
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using			
TITLE OF INVENTION: Thermostable FEN-1 Endonucleases			
NUMBER OF SEQUENCES: 190			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Medien & Carroll, LLP			
STREET: 220 Montgomery Street, Suite 2200			
CITY: San Francisco			
STATE: California			
COUNTRY: United States Of America			
Zip: 94104			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patent Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/757,653			
FILING DATE:			
CLASSIFICATION: 435			

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40, 027
; REFERENCE/DOCKET NUMBER: FOS- 02565
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 175:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1020
US-08-757-653-175

Query Match      11.0%; Score 162.8; DB 2; Length 1023;
Best Local Similarity 55.0%; Pred. No. 2,1e-35;
Matches 320; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

QY    146 CGATGAAGGAGCGCAAGAAGTTCGAGAGCTACTTTGTGGCCGAAATATGCCGTGCACGCCGCA 205
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     29 CAAGAAAAGAAATTGAGTTAGAAAAACCCTATACGGGAAAAAATAATGCATATGCACCTCTTA 88
QY    206 TGACATCTTACCACTTCCTGATATGTAAGTTGGAAAGCAGCGCATGSAACCTTCACAANTG 265
       ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     89 ATGCAATCTCCCAATTTTTTGTCACAAATVAGACAAGAAAGATGGAAGCTCCACTATATGATT 148
QY    266 AAGTGGTGTAAGTCATCACTAGTATTTGCCAAGAAATCTTCACCGSACAAATAGATTAATGC 325
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     149 CAAGGGGAAATTAACCTCCCACTPAAGCGGCTCTTTTACAGGCAATTAACCTAATGG 208
QY    326 AAGGGGAAATCAAGCCAGTTATGTTTGGATGGCAGAGCCTCTGATATGTAAGAAACAG 385
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     209 AGGCTGGAATTAACCTGCTGTATGTTTGGATGAGAACCTCCAGAAATTCAAAAAGAAAG 268
QY    386 AACTTCTAAATAGATCTCAAAAAAGATGATGCAACCAAGATCTGACTGAGCCACTAG 445
       ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     269 AGCTCGAAAAAAGAAAGAAAGACGACGAGAGAGCTGAAGAAAAGTGGAGACAAACACTTG 328
QY    446 AGGTAGAGNATTAAGATGCGATTGAAAAATTBAGCAAGAGCACTGTAAAGGTCAAGAGC 505
       ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     329 AAAAGAGGAGATGAGAGAGCAAGAAATATGGCCAAAAGACCAACCGATATATGAA 388
QY    506 AACCAACGAAGATTGTAAACGACTTTAAGACTATATGGGGTCCGTGCTGTAAGCAC 565
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     389 TGTCTATGAGAGATGCAAAAAAAGCTCTTAGAGCTATATGGGAATTCCTATPATTTAAACAC 448
QY    566 CTTCGAAAGCAGAACGAAATGTGCAAGCCCTTGGCATAAACGATTAAGCTGTCCGCTTGG 625
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     449 CTAGAGAGGAGAGAGCCCAAGCTGCATATATGACGCGCAAGGGGAGCGCTGATCAACGG 508
QY    626 CTTCGAGAGATTAAGAGTCCCTACTACTTTTGGGGGCTCCAAGGTTCCCTTGCTATTTAATGG 685
       ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     509 CTAGTCAGAAATTAAGATTCCCTACTTTTGGAGAGCTCCAAAGACTTGTATGAACCTTAACA 568
QY    686 ATCCAAAGTTCCAGAAAAATACCTGTGATGAGTAATTTGATGTTG 727
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     569 TAACGAGAAAAAGAAAGTTGCTCGGAAAAAAATGCTACGCTG 610

RESULT 11
US-08-823-516-78
; Sequence 78, Application US/08823516
; Patent No. 5994069
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Mast, Andrea L.
APPLICANT: Brow, Mary Ann D.
```

TITLE OF INVENTION: Detection of Nucleic Acids By Multiple  
 TITLE OF INVENTION: Sequential Invasive Cleavages  
 NUMBER OF SEQUENCES: 163  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Medien & Carroll, LLP  
 STREET: 220 Montgomery Street, Suite 2200  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States Of America  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/823,516  
 FILING DATE: 24-MAR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US97/01072  
 FILING DATE: 21-JAN-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/759,038  
 FILING DATE: 02-DEC-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/758,314  
 FILING DATE: 02-DEC-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/756,386  
 FILING DATE: 29-NOV-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/682,853  
 FILING DATE: 12-JUL-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/599,491  
 FILING DATE: 24-JAN-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ingolla, Diane E.  
 REGISTRATION NUMBER: 40,027  
 REFERENCE/DOCKET NUMBER: FORS-02736  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 705-8410  
 TELEFAX: (415) 397-8338  
 INFORMATION FOR SEQ. ID NO: 78:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1023 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..1020  
 US-08-823-516-78

Query Match	11.0%	Score 162.8	DB 2	Length 1023
Best Local Similarity	55.0%	Pred. No. 2.1e-35		
Matches	320	Conservative	0	Mismatches 262
				Indels 0
				Gaps 0
QY	146	CGATGAAGAGCAGAAAGTTGAGAGCTACTTTCGCGCAAAATGCGCGTGGAGCGCAGCA	205	
Db	29	CAGAAAAAGAAATGATGTGAAAAACCTATGCGGCAAAAAAATGCGCATGACGCTCTTAA	88	
QY	206	TGACCATCTACCGAGTTCCTGATAGTAGTACGTTGGAAAGACAGCGCATGAGAAACCTCCAAATG	265	
Db	89	ATGCAATCTCCCAATTTTTTGTCCACATATAAGACAAGAAAGATGGAATCTCCACTTATGAGTT	148	
QY	266	AAGCTGTGAAGTACCTAGTCATTTTTCAGAGAAATTTCAACCGGCACATAAGATTTACTGG	325	
Db	149	CAAAAGGTAGAAATTAACCTCCCACTAGTAAGGCGCTCTTTTACAGGACAAATTAACCTAATGG	208	
QY	326	AAGCGGGAATCAAGCCAGTTTATGTTTTTGATGGCAAGCTCTCGATATGGAAGAAACAG	385	

Db	209	AGCTTGAAATTAACAACTGTGTATGTTTGTATGAGACCTCCAGATTTCAAAAAAGAA	268
Qy	386	AACCTGCTAAAGATATCTCATAAAGAGATATGCACCAAAAGATCTGATGGGAGTAG	444.8
Db	269	AGCTCAAAAAAAGAGAGAGAGCGAGAGGGAGCTCAGAAAAGTGGAGAGACACTTG	322.8
Qy	446	AGGTAGAGATTAAGATGCGATTAATAAATTTGACCAAGAGACTGAAGCTCACAAAGC	505
Db	329	AAAAAGAGAGATAGCGAGAGCAAGAAAATATGCCCAAGACCAACAGGTAATATGAA	388
Qy	506	AACACACAGCAAGATGTATAACGACTATTAGACTATATGGGGGTCTCTGTTACAGGCAC	565
Db	389	TGCTCATCTCGAGGTGTCAAAAAACTTTAGAGCTTTATGGAAATCTCTATAGTTCACAGC	448
Qy	566	CTTCTGAGAGAGACAGATGTGCAGCCCTTTGCATAAACGATTAAGTGTCCGTGTTG	625
Db	449	CTAGCCAGGAGAGGCGCCACAGTGTGATATATGGCGGCAAAAGGAGACGTATGTCATCG	508
Qy	626	CTTCAGAGATTAAGACTCCCTTACTTATTGGGCTCCACGGTTCCTTGCTATTAAAGG	685
Db	509	CTAGTCAAGATATACGATTCCTCACTTTTGTGGAGCTCCAAACACTGTATTGAACCTTAACAA	566
Qy	686	ATCCAGATTCCAAGAAATATCCTGTGATGGAATTTGATGTTG	727
Db	569	TACACGAAAAAGCAAGTTCCTGGGAAAAATGTCTACGTGG	610

RESULT 12  
 US-08-759-038-114  
 : Sequence 114, Application US/08759038  
 : Patent No. 6090543  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Prudent, James R.  
 : APPLICANT: Hall, Jeff G.  
 : APPLICANT: Lyamichev, Victor I.  
 : APPLICANT: Biow, Mary Ann D.  
 : APPLICANT: Dahlberg, James E.  
 : TITLE OF INVENTION: Cleavage of Nucleic Acids  
 : NUMBER OF SEQUENCES: 114  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Medlen & Carroll, LLP  
 : STREET: 220 Montgomery Street, Suite 2200  
 : CITY: San Francisco  
 : STATE: California  
 : COUNTRY: United States of America  
 : ZIP: 94104  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent Release #1.0, Version #1.30  
 :  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/759,038  
 : FILING DATE: 02-DEC-1996  
 :  
 : CLASSIFICATION:  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/  
 : FILING DATE: 29-NOV-1996  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/682,853  
 : FILING DATE: 12-JUL-1996  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/599,491  
 : FILING DATE: 24-JAN-1996  
 :  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Ingolia, Diane E.  
 : REGISTRATION NUMBER: 40,027  
 : REFERENCE/DOCKET NUMBER: FORS-02574  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (415) 705-8410  
 : TELEFAX: (415) 397-8338  
 :  
 : INFORMATION FOR SEQ ID NO: 114:

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 1023 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1020
: US-08-759-038-114

```

```

Query Match      11.0%; Score 162.8; DB 3: Length 1023;
Best Local Similarity 55.0%; Pred. No. 2.1e-35;
Matches 320; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

```

```

QY 146 CGATGAGAGCAGCAAGTTCGAGAGCTACTTTCGCGCAAAATCGCGGTGACGCCAGCA 205
Db 29 CAAAGAAAGAAATTTGAGTTGAAAACCTATACGGGAAAAAATCGCAATCGACGCTCTTA 88
QY 206 TGACCATCTACCGAGTTCCTGATAGTGTGGAGAGCAGCATGGAACCTCCACAATG 265
Db 89 ATGCAATCTACCAATTTTGTTCACATTAAGACAGAAAGTGAAGTCCACTTATGATG 148
QY 266 AAGCTGTGAAGTCACTAGTATTGCAAGAAATGTTCAACCGACATATAGATTACTGG 325
Db 149 CAAAGGTAGAAATACCTCCACCTAAGCGGGCTCTTTTACAGACATATAACTATATG 208
QY 326 AAGGGGAATCAAGCCAGTTATGTTTGTGAGCAAGCTCCTGATATGAGAAAGCAAG 385
Db 209 AGGCTGGAATTAACCTGTATATGTTTGTATGAGACCTCCGATTCATCAAAAGAAAG 268
QY 386 AACTTGCTAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGGCAATG 445
Db 269 AGCTCGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 328
QY 446 AGTAGGAGATTAAGATGAGCATTTGAAAAATTTGAGCAAGAGCATGTAAGAGTCACAGGC 505
Db 329 AAAAAAGAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
QY 506 AACCAACAGAGATTTGAAACGACTTATAGACTTATGAGGGGTTCTGTTGTAGAGGCAC 565
Db 389 TGCTATCGAGAGATGCAAAAAAACTCTTAGAGCTTATGGGAATCTCTATAGTTCAGAGC 448
QY 566 CTCTGAGAGAGAGAGAGAGATGTCAGACCTTTGCAATTAACATTAAGTGTGCTGTG 625
Db 449 CTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508
QY 626 CTTCAGAGATTAAGAGACTCCCTTACTTGTGGGCTCCAGGTTCTCTCTGCTATTTATG 685
Db 509 CTAGTCAAGATTAAGATGATTTCTTACTTGTGGAGCTCCAGAGCTGTATTAAGAACTTA 568
QY 686 ATCCAAGTTCAGAAATATACCTGTGATGGAATTTGATGCTG 727
Db 569 TAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610

```

## RESULT 13

```

US-08-758-314-114
: Sequence 114, Application US/08758314
: Patent No. 6090606

```

```

: GENERAL INFORMATION:
: APPLICANT: Kaiser, Michael W.
: APPLICANT: Lyamichiev, Victor I.
: APPLICANT: Lyamichiev, Natasha
: TITLE OF INVENTION: Improved Cleavage Agents
: NUMBER OF SEQUENCES: 134
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Medlen & Carroll, LLP
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: COUNTRY: United States of America
: ZIP: 94104

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/758,314
: FILING DATE: 02-DEC-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/
: FILING DATE: 29-NOV-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/682,853
: FILING DATE: 12-JUL-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/599,491
: FILING DATE: 24-JAN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Ingolia, Diane E.
: REGISTRATION NUMBER: 40,027
: REFERENCE/DOCKET NUMBER: FORS-02575
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 705-8410
: TELEFAX: (415) 397-8338
: INFORMATION FOR SEQ. ID NO.: 114:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1023 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1020
: US-08-758-314-114

```

```

Query Match      11.0%; Score 162.8; DB 3: Length 1023;
Best Local Similarity 55.0%; Pred. No. 2.1e-35;
Matches 320; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

```

```

QY 146 CGATGAGAGCAGCAAGTTCGAGAGCTACTTCGCGCAAAATCGCGGTGACGCCAGCA 205
Db 29 CAAAGAAAGAAATTTGAGTTGAAAACCTATACGGGAAAAAATCGCAATCGACGCTCTTA 88
QY 206 TGACCATCTACCAATTTTGTTCACATTAAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 265
Db 89 ATGCAATCTACCAATTTTGTTCACATTAAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 148
QY 266 AAGCTGTGAAGTCACTAGTATTGCAAGAAATGTTCAACCGGACATTAAGATTACTGG 325
Db 149 CAAAGGTAGAAATACCTCCACCTAAGCGGGCTCTTTTACAGAGCAATTAACCTTAATG 208
QY 326 AAGCGGAATCAAGCAGATTTATGTTTGTATGCAAGGCTCCTGATTAAGAAAGCAAG 385
Db 209 AGGCTGGAATTAACCTGCTATGTTTGTATGATGAGAGAACTCCAGAAATTCAGAAAGAA 268
QY 386 AACTTGCTAAAGATTAAGTCTCAAAAAGAGATGCAACCAAGATGTGACTAGAGGAGAG 445
Db 269 AGCTCGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 328
QY 446 AGTAGAGATTAAGATGCGATTAAGAAATTTGAGCAAGAGAGAGAGAGAGAGAGAGAGAG 505
Db 329 AAAAAAGAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
QY 506 AACCAACAGAGATTTGAAACGACTTATAGACTTATGAGGGGTTCTGTTGTAGAGGCAC 565
Db 389 TGCTATCGAGAGATGCAAAAAAACTTTAGAGCTTATGAGGAATTCCTTATGTTCAAGAC 448
QY 566 CTCTGAGAGAGAGAGAGAGATGTCAGACCTTTGCAATTAACATTAAGTGTGCTGCTTG 625
Db 449 CTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508

```





Tue May 27 16:02:32 2003

us-09-805-311-7.std.rni

Page 13

Oy	626	CTTCGAGATTAAGACACCCTTACTCCTTTGGGGGTCCACGCSTTCCCTCGCATTTAATCG	885
Db	509	CGAGTCAGATTAAATGATTCACACTCTTGCTGGTCCCAAGTCATATTAGCAATCTGCMA	568
Oy	686	ATCCAGATTCCACAAGAATACTGATGATGCATTTGATGTGG	727
Dd	569	TTCACGGAAAAAGAAAGATGCTGGGAAGATGTTTACGTGG	610

Search completed: November 5, 2002, 11:03:14  
Job time : 46.5823 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw.model

Run on: November 5, 2002, 10:28:52 : Search time 1264.61 Seconds  
(without alignments)  
15774.429 Million cell updates/sec

Title: US-09-805-311-7

Perfect score: 1478  
Sequence: 1 cgaaccacgcgtccgggaaa.....aaaaaaaaaaaaaaaa 1478

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estrpl:\*  
7: em\_estro:\*  
8: em\_etc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	793	53.7	901	10	BG837708 zml0_01f0
2	570	38.6	586	9	AI881599 606068G09
3	538.2	36.4	553	9	AM562789 660065H06
4	534.2	36.1	554	10	BE639421 946033A02
5	530	35.9	550	10	BE639422 946033A02
6	497.8	33.7	532	9	AM000375 614014D03
7	471.4	31.9	474	9	AM559173 660065H06
8	464	31.4	468	9	AV913663 AV913663
9	454.4	30.7	467	9	BE186786 946012C08
10	454	30.6	470	9	AI861468 614014D03
11	451.8	30.6	475	9	AM562517 660065H06
12	435	29.4	553	9	AI834484 606068G09
13	417.2	28.2	733	9	AM774700 EST333851
14	408.4	27.6	539	10	BG263233 WHE2339_F
15	399	27.0	414	9	AM288784 707010C02
16	393.2	26.6	456	9	AI065689 ag91f12.x
17	375.6	25.4	632	9	AM585913 EST317536

18	370.4	25.1	754	10	BE416584	BE416584	MUC009.B1
19	369.6	25.0	820	10	BG414505	BG414505	HVSMK000
20	369.4	25.0	609	10	BG096545	BG096545	EST461064
21	361.4	24.5	645	9	AM691681	AM691681	NE042B095
22	359.8	24.3	363	9	AM562518	AM562518	660065H06
23	348	23.5	680	10	BE417817	BE417817	MUG024.H1
24	347.2	23.5	463	9	AV918966	AV918966	AV918966
25	345.2	23.4	454	9	AV915644	AV915644	AV915644
26	337	22.1	539	10	BF098320	BF098320	EST428761
27	316	21.4	592	9	AT780966	AT780966	EST261845
28	314	21.2	470	9	AM288831	AM288831	707010E11
29	311.6	21.1	579	9	AV833393	AV833393	AV833393
30	310	21.0	376	9	AI065546	AI065546	ag98e02.x
31	306.4	20.7	526	10	B1419720	B1419720	LJNESt479
32	295.8	20.0	695	9	BE052579	BE052579	CA_Pa000
33	271.4	18.4	549	9	AM696665	AM696665	NP109F01S
34	267.6	18.1	942	9	AL560007	AL560007	AL560007
35	259.4	17.6	903	9	AL531350	AL531350	AL531350
36	257.6	17.4	851	9	AL519300	AL519300	AL519300
37	243.8	16.5	689	10	BE311755	BE311755	601143566
38	240.6	16.3	769	10	BG574950	BG574950	602598038
39	238	16.1	660	10	Bj011893	Bj011893	AL560395
40	238	16.1	806	9	AL560395	AL560395	602435150
41	232.8	15.8	762	10	BG337603	BG337603	rp09g07.Y
42	232	15.7	704	9	AM783795	AM783795	603073821
43	229.8	15.5	866	10	B1827898	B1827898	603073821
44	229	15.5	812	10	BM015629	BM015629	603041864
45	228.6	15.5	963	12	CNS06WXR	AL419029	T3 end of

## ALIGNMENTS

RESULT 1  
LOCUS BG837708 901 bp mRNA linear EST 25-MAY-2001  
DEFINITION zml0\_01f08\_A zml0\_AAFC\_ECORC\_Fusarium\_graminearum\_corn\_silk zea  
mays cDNA clone zml0\_01f08, mRNA sequence.  
ACCESSION BG837708.1 GI:14204031  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Zea mays.  
Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 901)  
Harris,L.J., Balcerzak,M., Allard,S., Saparno,A., Couroux,P., De  
Moors,R., Hattori,J.I., Ouellet,T., Robert,L.S., Singh,J.A., Sprott  
D. and Linker,N.A.

## REFERENCE AUTHORS

Expressed Sequence Tags from Maize Silk Six Hours After Silk  
Channel Inoculation with Fusarium graminearum  
Unpublished (2001)  
Contact: Harris, Linda J.  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-food Canada  
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
CANADA  
Tel: (613) 759-1314  
Fax: (613) 759-6566  
Email: harris@em.agr.ca.

## JOURNAL COMMENT

## FEATURES source

Location/Qualifiers  
1..901  
/organism="Zea mays"  
/cultivar="CO388"  
/db\_xref="taxon:4577"  
/clone="zml0\_01f08"  
/clone\_lib="zml0\_AAFC\_ECORC\_Fusarium\_graminearum\_corn\_silk"  
/tissue\_type="Silk"  
/dev\_stage="4-5 days post-silk emergence"  
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site\_1: EcoRI;  
Site\_2: XhoI; Field-grown corn was silk channel-inoculated

in the morning (-10 am) with 1 ml of a macroconidial suspension (500,000 spores/ml) of *Fusarium graminearum* and silk channels were collected and immediately frozen in liquid nitrogen 6 hours later. RNA was extracted from silk tissue between 1 cm below and above the inoculation point in the silk channel, RNA from five silk channels was pooled."

BASE COUNT 267 a 174 c 208 g 249 t 3 others  
ORIGIN

Query Match 53.7%; Score 793; DB 10; Length 901;  
Best Local Similarity 98.6%; Pred. No. 3.5e-124;

Matches 829; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

QY 617 TCGCTGTTGCTTGAAGATAGAGATCCCTTACTTTGGGGCTCCAGGTTCTTCGTC 676  
Db 1 TCGCTGTTGCTTGAAGATAGAGATCCCTTACTTTGGGGCTCCAGGTTCTTCGTC 60  
QY 677 ATTTAATGATCCAGTTCCTCAAGAAATACCTGTGATGATTTGTGTCGCAAGTTT 736  
Db 61 ATTTAATGATCCAGTTCCTCAAGAAATACCTGTGATGATTTGTGTCGCAAGTTT 120  
QY 737 TGGAGAGCTTGAACCTCCAGTTCATTTGATGATCTGTGATGTC 796  
Db 121 TGGAGAGCTTGAACCTCCAGTTCATTTGATGATCTGTGATGTC 180  
QY 797 ACTATGTTGATGATCAAGATATGGGGGCAACACCTGGAACCTATTCGTCAAC 856  
Db 181 ACTATGTTGATGATCAAGATATGGGGGCAACACCTGGAACCTATTCGTCAAC 240  
QY 857 ATGGGTCATAGAAGACTTGTGAGATCTTAAATAAGACAGATATCAATTC -TGAG 915  
Db 241 ATGGGTCATAGAAGACTTGTGAGATCTTAAATAAGACAGATATCAATTC -TGAG 300  
QY 916 GACGAGCTTAC -AAGAACTGAGCTTGTCAAGAGAGCCATATGCTTTGGATAT 974  
Db 301 GACGAGCTTAC -AAGAACTGAGCTTGTCAAGAGAGCCATATGCTTTGGATAT 360  
QY 975 TCCGAGCTTAAATGAGCTGACCTGATGAGAGGGCTCATTAAGTTCTCTGTTAAAGA 1034  
Db 361 TCCGAGCTTAAATGAGCTGACCTGATGAGAGGGCTCATTAAGTTCTCTGTTAAAGA 420  
QY 1035 TAAATGTTCAATGAGATGGGTGACCAAGGCTAGAGAGATCAATTCGCCAAGA 1094  
Db 421 TAAATGTTCAATGAGATGGGTGACCAAGGCTAGAGAGATCAATTCGCCAAGA 480  
QY 1095 TAAATGTTCAATGAGATGGGTGACCAAGGCTAGAGAGATCAATTCGCCAAGA 1154  
Db 481 TAAATGTTCAATGAGATGGGTGACCAAGGCTAGAGAGATCAATTCGCCAAGA 540  
QY 1155 GCTAAAGCGAAGAGACTCGGATTAACACAGCAGGAGCTGGCAACAAAGCAAA 1214  
Db 541 GCTAAAGCGAAGAGACTCGGATTAACACAGCAGGAGCTGGCAACAAAGCAAA 600  
QY 1215 GCGTGTGGAAGAAGAAATATCTTGATGCTTGTATGATCACTAGCAAGCA 1274  
Db 601 GCGTGTGGAAGAAGAAATATCTTGATGCTTGTATGATCACTAGCAAGCA 660  
QY 1275 GCGGTGATGATGATCTTCCCTAGATATTTACTCCCTGTTTACTCAACCTTTGG 1334  
Db 661 GCGGTGATGATGATCTTCCCTAGATATTTACTCCCTGTTTACTCAACCTTTGG 720  
QY 1335 TGAAGTTTCCCATGTTTAAAGCTGGGTAAAGTATGTTGTTGAAGATTTGTTA 1394  
Db 721 TGAAGTTTCCCATGTTTAAAGCTGGGTAAAGTATGTTGTTGTTGAAGATTTGTTA 780  
QY 1395 CCAAGTAAACAACCTTATGCTG -TTTTTACTTCTTGTCTTGAAGATATGATCCAG 1453  
Db 781 CCAAGTAAACAACCTTATGCTGCTTTTACTTCTTGTCTTGAAGATATGATCCAG 840  
QY 1454 T 1454  
Db 841 T 841

RESULT 2  
LOCUS A1881599 586 bp mRNA linear EST 02-FEB-2000  
DEFINITION 606068609.y1 606 - Ear tissue cDNA library from Schmidt lab Zea  
mays cDNA, mRNA sequence.  
ACCESSION A1881599  
VERSION A1881599.1 GI:5566733  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC  
clade; Panicoidae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 586)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY  
JOURNAL Unpublished (1995)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Place: 606068 row: 6 column: 09.  
FEATURES  
source  
1..586  
/organism="Zea mays"  
/cultivar="Ohio43"  
/db\_xref="taxon:4577"  
/clone\_lib="606 - Ear tissue cDNA library from Schmidt lab"  
/tissue\_type="mixed"  
/dev\_stage="ear length from 0.5 cm - 2.0 cm"  
/lab\_host="XLOLR (Stratagene)"  
/note="organ: immature ear; Vector: pBR-CMV; Site\_1: EcoRI  
Site\_2: XhoI; Mixed ear tissue cDNA library from Schmidt lab"  
BASE COUNT 186 a 132 c 159 g 109 t  
ORIGIN  
Query Match 38.6%; Score 570; DB 9; Length 586;  
Best Local Similarity 100.0%; Pred. No. 1.4e-86;  
Matches 570; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 16 GGAATAGCTCGGGGTCGGGCTTCTTGGCCACTCCGGCTCAGCGCGCCGCAACC 75  
Db 17 GGAATAGCTCGGGGTCGGGCTTCTTGGCCACTCCGGCTCAGCGCGCCGCAACC 76  
QY 76 GCCACAGCCCGCCGACAGATGGGCTTGTGACGAACCTGCTGGCGACAT 135  
Db 77 GCCACAGCCCGCCGACAGATGGGCTTGTGACGAACCTGCTGGCGACAT 136  
QY 136 GCCGCCAAGCGGTGAAGAGCAGAGTTCGAGACTCTTCCGCGCAAAATCGCCGTC 195  
Db 137 GCCGCCAAGCGGTGAAGAGCAGAGTTCGAGACTCTTCCGCGCAAAATCGCCGTC 196  
QY 196 GAGCGCAGATGAGATCTACAGTTCTGATATGATTTGAAGAGCAGCATGAAAT 255  
Db 197 GAGCGCAGATGAGATCTACAGTTCTGATATGATTTGAAGAGCAGCATGAAAT 256  
QY 256 CTCGCAATGAACCTGCTGAAGCACTAGTCTTGTGCAAGCAAGTTCAACCGGACATA 315  
Db 257 CTCGCAATGAACCTGCTGAAGCACTAGTCTTGTGCAAGCAAGTTCAACCGGACATA 316  
QY 316 AGATTACTGGAAGCGGAATCAAGCGATTATGTTTTGATGAGAGCTCTGTATATG 375  
Db 317 AGATTACTGGAAGCGGAATCAAGCGATTATGTTTTGATGAGAGCTCTGTATATG 376  
QY 376 AAGAACAGAACTTCTTAAAGATACCTCAAAAAGATGATGCAACGAAGATCTGACT 435

```

Db 377 AAGAAACAGACTTGTCTAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTACT 436
Oy 436 GAGCAGTAGAGTAGAGATGAATGCGATGTGAAAAATGAGCAGAGAGACTGTAAG 495
Db 437 GAGCAGTAGAGTAGAGATGAATGCGATGTGAAAAATGAGCAGAGAGACTGTAAG 496
Oy 496 GTCAACAGGCAACACAAAGAGATTGTAACGACTATTAGACTTTATGGGGCTTCGTGTT 555
Db 497 GTCAACAGGCAACACAAAGAGATTGTAACGACTATTAGACTTTATGGGGCTTCGTGTT 556
Oy 556 GTAGAGGACCTCTGGAAGCAGAGCAGAA 585
Db 557 GTAGAGGACCTCTGGAAGCAGAGCAGAA 586

RESULT 3
AM562789 553 bp mRNA linear EST 10-MAR-2000
LOCUS 660065H06.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AM562789
VERSION AM562789.1 GI:7216667
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 553)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.
FEATURES
source
1..553
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 171 a 111 c 129 g 142 t
ORIGIN
Query Match 36.4%; Score 538.2; DB 9; Length 553;
Best Local Similarity 99.4%; Pred. No. 3.3e-81;
Matches 540; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 915 GCATGCGCTTACCAAGAGCTCGACGCTGTGTCAGAGAGCCTATGTCACCTTGGATAT 974
Db 4 GGACTGGCTTACCAAGAGCTCGACGCTGTGTCAGAGAGCCTATGTCACCTTGGATAT 63
Oy 975 TCCTAGCTTAATAATGAGCTGCACCTGATGAGAGGGTCTCATAGTTCTGTGTAAGA 1034
Db 64 TCCTAGCTTAATAATGAGCTGCACCTGATGAGAGGGTCTCATAGTTCTGTGTAAGA 123
Oy 1035 TAATGCTTCAATGAAGATCGGGTGACAAAGGCCATAGAGAAGATCAATCTGCCAAGA 1094
Db 124 TAATGCTTCAATGAAGATCGGGTGACAAAGGCCATAGAGAAGATCAATCTGCCAAGA 183

```

```

Oy 1095 TAATGCTGCAAGAGAAAGCTGATGCTTTTCAAGCCAACTGCCACATCAGACCC 1154
Db 184 TAATGCTGCAAGAGAAAGCTGATGCTTTTCAAGCCAACTGCCACATCAGACCC 243
Oy 1155 GCTAAACGGAAGAGACTTGGATTAACCAAGCAGGACGCTGCCAACAAGAAAAA 1214
Db 244 GCTAAACGGAAGAGACTTGGATTAACCAAGCAGGACGCTGCCAACAAGAAAAA 303
Oy 1215 GCGTGTGTAAGAAAGAAATATCTTGATGCTGTGATGATCACTACGATCAGAA 1274
Db 304 GCGTGTGTAAGAAAGAAATATCTTGATGCTGTGATGATCACTACGATCAGAA 363
Oy 1275 GCGTGTGATGATCACTGCTGCTAGATTAATTAACCTCCGTTTAACTCAGACCTTTGG 1334
Db 364 GCGTGTGATGATCACTGCTGCTAGATTAATTAACCTCCGTTTAACTCAGACCTTTGG 423
Oy 1335 TGAAGATTTGCCATGTTTCAGCTGGGTAAGTATGTTGTGTAAGAGATTGGTGA 1394
Db 424 TGAAGATTTGCCATGTTTCAGCTGGGTAAGTATGTTGTGTAAGAGATTGGTGA 483
Oy 1395 CCAAGTACAAACTATCGCTGTTTAACTTCTGTGCTTGAAGTATGATGCGAGT 1454
Db 484 CCAAGTACAAACTATCGCTGTTTAACTTCTGTGCTTGAAGTATGATGCGAGT 543
Oy 1455 AAA 1457
Db 544 ACA 546

RESULT 4
BE639421/c 554 bp mRNA linear EST 30-AUG-2000
LOCUS 946033A02.x2 946 - tassels primordium prepared by Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION BE639421
VERSION BE639421.1 GI:9952838
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 554)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946033 row: A column: 02.
FEATURES
source
1..554
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassel primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybridZAP; Site_1: EcoRI;
Site_2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybridZAP. Sample insert size range was 350 bp
to 3 kb with a 1 kb average."
BASE COUNT 145 a 131 c 110 g 168 t

```



ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 532)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 614014 row: D column: 03.

FEATURES  
source location/Qualifiers  
1..532  
/organism="Zea mays"  
/cultivar="W23"  
/db\_xref="taxon:4577"  
/clone\_lib="614 - root cDNA library from Walbot Lab"  
/tissue\_type="root"  
/dev\_stage="3-4 days old"  
/lab\_host="X10LR"  
/note="Organ: root; Vector: pBluescriptII SK+; Site\_1: EcoRI; Site\_2: XhoI; 3-4 days old root tissue from Walbot Lab (LM)"

BASE COUNT 173 a 103 c 129 g 127 t

Query Match 33.7%; Score 497.8; DB 9; Length 532;  
Best Local Similarity 98.3%; Pred. No. 2.2e-74;  
Matches 524; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 678 TTTATGATCCAAAGTTCCAGAAATACCTGATGAGATTTGATGTTGCCAAGTTT 737  
Db 1 TTTATGATCCAAAGTTTCAAGAAATACCTGATGAGATTTGATGTTGCCAAGTTT 60  
QY 738 GGAGAGCTTGAATCACTACCATGAGCAGTTGATTTGTGCATCCTGTGATGTA 797  
Db 61 GGAGAGCTTGAATCACTACCATGAGCAGTTGATTTGTGCATCCTGTGATGTA 120  
QY 798 CTATTGATGATCAAAAGTATCGGGGGCAACAGCTCTCAACTTATTCGTCAACA 857  
Db 121 CTATTGATGATCAAAAGTATCGGGGGCAACAGCTCTCAACTTATTCGTCAACA 180  
QY 858 TGGTCCATAGAAAGCATCTTGGAGATCTTAATAAGACA-GATATCAATTCCTGAG 916  
Db 181 TGGTCCATAGAAAGCATCTTGGAGATCTTAATAAGAGATATCAATTCCTGAG 240  
QY 917 ACTGGCCTTACCAAGAGCTGCAGCCTTGTCAAGAGCCTATATGCACTTTGGATTC 976  
Db 241 ACTGGCCTTACCAAGAGCTGCAGCCTTGTCAAGAGCCTATATGCACTTTGGATTC 299  
QY 977 CTGAGCTAAATGAGTGCACCTGATGAGAGGGTCTCAATGATTTCTGTAAGATA 1036  
Db 300 CTGAGCTAAATGAGTGCACCTGATGAGAGGGTCTCAATGATTTCTGTAAGATA 359  
QY 1037 ATGTTTCAATGAGATCGGGTGACAAAGGCCATAGAAAGTCAAAATGTCACAGATA 1096  
Db 360 ATGTTTCAATGAGATCGGGTGACAAAGGCCATAGAAAGTCAAAATGTCACAGATA 419  
QY 1097 AATGCTGCAAGAAAGTCTGAGTCTTTTTCAGGCCAATGTCACACATGAGACCGC 1156  
Db 420 AATGCTGCAAGAAAGTCTGAGTCTTTTTCAGGCCAATGTCACACATGAGACCGC 479  
QY 1157 TAAACGGAAGAGACTTGGATTAACCAAGCAAGGAGCTGGAGCAAGAAA 1209  
Db 480 TAAACGGAAGAGACTTGGATTAACCAAGCAAGGAGCTGGAGCAAGAAA 532

RESULT 7  
AM559173/c 474 bp mRNA linear EST 07-MAR-2000  
LOCUS 660065H06.xl 660 - Mixed stages of anther and pollen Zea mays cDNA,  
DEFINITION mRNA sequence.  
ACCESSION AM559173  
VERSION AM559173.1 GI:7204640  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 474)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 660065 row: H column: 06.

FEATURES  
source location/Qualifiers  
1..474  
/organism="Zea mays"  
/cultivar="Ohio43"  
/db\_xref="taxon:4577"  
/clone\_lib="660 - Mixed stages of anther and pollen"  
/tissue\_type="whole premeiotic anthers to pollen shed"  
/dev\_stage="premeiotic anthers to pollen shed"  
/lab\_host="X10LR"  
/note="Organ: anthers; Vector: Lambda Zap; Site\_1: EcoRI; Site\_2: XhoI; Anther and pollen cDNA library. Directionally sequenced with 5' end at the EcoRI site. Created by Amie Franklin."

BASE COUNT 113 a 114 c 92 g 154 t 1 others

Query Match 31.9%; Score 471.4; DB 9; Length 474;  
Best Local Similarity 99.6%; Pred. No. 6.2e-70;  
Matches 472; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 932 AAGCTGACGCTTGTTCAGAGAGCCATAATGTCACCTTGGATTTCTGAGCTAAATGGA 991  
Db 474 AAGCTGACGCTTGTTCAGAGAGCCATAATGTCACCTTGGATTTCTGAGCTAAATGGA 415  
QY 992 CTGCACCTGATGAGAGGGTCTCATAGATTCCTGTTAAAGATTAATGTTCAATGAAG 1051  
Db 414 CTGCACCTGATGAGAGGGTCTCATAGATTTCTGTTAAAGATTAATGTTCAATGAAG 355  
QY 1052 ATCGGGTGACAAAGCCATAGAGAAATCAAAATCTGCCAAGATTAATCGTCGACGAAA 1111  
Db 354 ATCGGGTGACAAAGCCATAGAGAAATCAAAATCTGCCAAGATTAATCGTCGACGAAA 295  
QY 1112 GACTCGAGTCTTTTTCAGGCCACATCGCCACCATCGACCCCTAAAGAGAGAGA 1171  
Db 294 GACTCGAGTCTTTTTCAGGCCACATCGCCACCATCGACCCCTAAAGAGAGAGA 235  
QY 1172 CTTCGGATAAACAAGCAAGCAGCTGCGAACAGAAAACAAGCTGTGTGAAAGAGA 1231  
Db 234 CTTCGGATAAACAAGCAAGCAGCTGCGAACAGAAAACAAGCTGTGTGAAAGAGA 175  
QY 1232 AATAATCTTGATGATGTTGATGATCAACTAGACACTAGAAACAGACGCTGACATGACT 1291  
Db 174 AATAATCTTGATGATGTTGATGATGATCAACTAGACACTAGAAACAGACGCTGACATGACT 115  
QY 1292 TCGCCTAGATTAATTTACCTCCCTGTTTAACTCAACACTTGTGTGAAAGTTGCCATGT 1351  
Db 114 TCGCCTAGATTAATTTACCTCCCTGTTTAACTCAACACTTGTGTGAAAGTTGCCATGT 55

```

QY 1352 TTCAAGCTGGGGTAACTTGTCTTTCAAGAGATTTGCTGACCACTAACAA 1405
Db 54 TTCAAGCTGGGGTAACTTGTCTTTCAAGAGATTTGCTGACCACTAACAA 1

RESULT 8
AV913663 648 bp mRNA linear EST 18-JAN-2002
LOCUS AV913663 K. Sato unpublished cDNA library, cv. Haruna Nijo
DEFINITION germination shoots Hordeum vulgare subsp. vulgare cDNA clone
          bags2n17 5', mRNA sequence.
ACCESSION AV913663
VERSION AV913663.1 GI:18209440
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare.
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
          ; Triticeae; Hordeum.
          1 (bases 1 to 648)
REFERENCE Sato, K., Saito, D. and Takeda, K.
AUTHORS Barley EST sequencing project in NIG and Okayama Univ
TITLE Unpublished (2002)
JOURNAL Contact: Tadasu Shin-i
          Center for Genetic Resource Information
          National Institute of Genetics
          1111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshin@genes.nig.ac.jp.
FEATURES
  source
    1..648
    /organism="Hordeum vulgare subsp. vulgare"
    /cultivar="Haruna Nijo"
    /db_xref="taxon:112509"
    /clone="bags2n17"
    /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
    Nijo germination shoots"
    /issue_type="shoots"
    /dev_stage="germination"

BASE COUNT 174 a 158 c 185 g 131 t
ORIGIN
Query Match 31.4%; Score 464; DB 9; Length 648;
Best Local Similarity 82.3%; Pred. No. 1e-68;
Matches 533; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```

```

Db 361 AAGAGAGAGACAGCTTTTAAAAAGACACCGAAGAGATTAAGCAAGAGACTACG 420
QY 436 GAGCAGATAGAGATTAAGATTAAGATGCGATTGAAAAATTGACAGAGACTGTAAG 495
Db 421 AAGCAGATAGAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 480
QY 496 GTACAGAGAGAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 555
Db 481 GTACAGAGAGAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 540
QY 556 GTAGAGAGAGAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 615
Db 541 GTAGAGAGAGAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 600
QY 616 TTGCTGTTGCTTCAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663
Db 601 TATGCTGTTGCTTCAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648

```

```

RESULT 9
BE186786/c 467 bp mRNA linear EST 22-JUN-2000
LOCUS BE186786
DEFINITION 946012C08.X1 946 - tassal primordium prepared by Schmidt lab Zea
          mays cDNA, mRNA sequence.
ACCESSION BE186786
VERSION BE186786.1 GI:8665970
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
          clade; Panicoidae; Andropogoneae; Zea.
          1 (bases 1 to 467)
REFERENCE Walbot, V.
AUTHORS Maize ESTs from various cDNA libraries sequenced at Stanford
TITLE University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
          Department of Biological Sciences
          Stanford University
          855 California Ave, Palo Alto, CA 94304, USA
          Tel: 650 723 2227
          Fax: 650 725 8221
          Email: walbot@stanford.edu
          Plate: 946012 row: C column: 08.
FEATURES
  source
    1..467
    /organism="Zea mays"
    /cultivar="OH43"
    /db_xref="taxon:4577"
    /clone="946 - tassal primordium prepared by Schmidt
    lab"
    /clone_lib="946 - tassal primordium prepared by Schmidt
    lab"
    /issue_type="tassels"
    /dev_stage="just after the transition from vegetative to
    inflorescence development"
    /lab_host="XLOLR"
    /note="organ: tassels; Vector: HybriZAP; Site_1: EcoRI;
    Site_2: XhoI; George Chnuck dissected immature tassels
    between 1mm and 3mm. Sharon Stanfield prepared the cDNA
    library in HybriZAP. Sample insert size range was 350 bp
    to 3 kb with a 1 kb average."

BASE COUNT 108 a 110 c 96 g 153 t
ORIGIN
Query Match 30.7%; Score 454.4; DB 9; Length 467;
Best Local Similarity 99.6%; Pred. No. 4.6e-67;
Matches 466; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

QY 898 AGATATCAATTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
Db 467 AGATATCAATTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 408

```





BASE COUNT 124 a 110 c 94 g 147 t  
 ORIGIN

Query Match 30.6%; Score 451.8; DB 9; Length 475;  
 Best Local Similarity 99.4%; Pred. No. 1.2e-66;  
 Matches 464; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 991 ACTGACCTGATGAGGGGTCTCATAGTTCTCGTAAACATATGTTTCATGAA 1050  
 |||||||  
 DB 475 ACTGACCTGATGAGGGGTCTCATAGTTCTCGTAAACATATGTTTCATGAA 416  
 OY 1051 GATCGGGTGACCAAGGCCATAGAGAGATCAATCTGCCAAGTAATCTGCGCAAGSA 1110  
 |||||||  
 DB 415 GATCGGGTGACCAAGGCCATAGAGAGATCAATCTGCCAAGTAATCTGCGCAAGSA 356  
 OY 1111 AGACTCGAGTCTTTTCAAGCCACATGCGACACATACACCGCTTAAACGGAGAGAG 1170  
 |||||||  
 DB 355 AGACTCGAGTCTTTTCAAGCCACATGCGACACATACACCGCTTAAACGGAGAGAG 296  
 OY 1171 ACTTCGGATAAACACAGCAGCAGTGGCAGACAGAAACAAAGGCTGTGGAAGAAG 1230  
 |||||||  
 DB 295 ACTTCGGATAAACACAGCAGCAGTGGCAGACAGAAACAAAGGCTGTGGAAGAAG 236  
 OY 1231 AATATATCTTGATGCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1290  
 |||||||  
 DB 235 AATATATCTTGATGCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 176  
 OY 1291 TTTCGCTGATTTTAACTCCCTGTTTAACTCAGACCTTTGTGAAGTTTGGCCATG 1350  
 |||||||  
 DB 125 TTTCGCTGATTTTAACTCCCTGTTTAACTCAGACCTTTGTGAAGTTTGGCCATG 116  
 OY 1351 TTTCAGCTGGGTAAGTGTGTTTGAAGATGTTGATGATGATGATGATGATGATGATGAT 1410  
 |||||||  
 DB 115 TTTCAGCTGGGTAAGTGTGTTTGAAGATGTTGATGATGATGATGATGATGATGATGAT 57  
 OY 1411 ATCCGCTGTTTAACTCTTCTGCTTGAAGATGATGATGATGATGATGATGATGATGAT 1457  
 |||||||  
 DB 56 ATCCGCTGTTTAACTCTTCTGCTTGAAGATGATGATGATGATGATGATGATGATGAT 10

RESULT 12  
 A1834484/c 553 bp mRNA linear EST 02-FEB-2000  
 LOCUS 60606609.xl 606 - Ear tissue cDNA library from Schmidt lab Zea  
 DEFINITION mays cDNA, mRNA sequence.  
 ACCESSION A1834484  
 VERSION A1834484.1 GI:5468693  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 553)  
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
 University  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Walbot, V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 606068 row: G column: 09.

# FEATURES

source  
 1. 553  
 /organism="Zea mays"  
 /culturivar="Oh1043"  
 /db\_xref="taxon:4577"  
 /clone\_lhb="606 - Ear tissue cDNA library from Schmidt  
 lab"

BASE COUNT 140 a 112 c 106 g 195 t  
 ORIGIN  
 /tissue\_type="mixed"  
 /dev\_stage="ear length from 0.5 cm - 2.0 cm"  
 /lab\_host="XLOLR (Stratagene)"  
 /note="Organ: Immature ear; Vector: pBK-CMV; Site\_1: EcoRI  
 ; Site\_2: XhoI; Mixed ear tissue cDNA library from Schmidt  
 lab"

Query Match 29.4%; Score 435; DB 9; Length 553;  
 Best Local Similarity 97.8%; Pred. No. 8.1e-64;  
 Matches 441; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1007 AGGTCATCAATGTTCCGTGTAAGATATGTTTCATGATGATGATGATGATGATGATGATGAT 1066  
 |||||||  
 DB 553 AGGTCATCAATGTTCCGTGTAAGATATGTTTCATGATGATGATGATGATGATGATGATGAT 494  
 OY 1067 CCATAGGAAGATCAATCTGCCAAGATTAATCTGCGCAAGAGAGTCCGATTT 1126  
 |||||||  
 DB 493 CCATAGGAAGATCAATCTGCCAAGATTAATCTGCGCAAGAGAGTCCGATTT 434  
 OY 1127 TCAAGCCAACTGCCACACATCAGCACCCCTTAAACGGAAGAGACTTCGATTAACAA 1186  
 |||||||  
 DB 433 TCAAGCCAACTGCCACACATCAGCACCCCTTAAACGGAAGAGACTTCGATTAACAA 374  
 OY 1187 GCAAGGAGCTGGGACACAGAAACAAAGGCTGTGGAAGAAGAAATATCTTGATGC 1246  
 |||||||  
 DB 373 GCAAGGAGCTGGGACACAGAAACAAAGGCTGTGGAAGAAGAAATATCTTGATGC 314  
 OY 1247 TTGATGTACACTACACATCAGAAACAGCAGGCTGATGATGATGATGATGATGATGATGAT 1306  
 |||||||  
 DB 313 TTGATGTACACTACACATCAGAAACAGCAGGCTGATGATGATGATGATGATGATGATGAT 254  
 OY 1307 AACCTCCGTTTAACTCAGACCTTTGGTGAAGTTTGGCCATGTTTCAAGTGGGGTAA 1366  
 |||||||  
 DB 253 AACCTCCGTTTAACTCAGACCTTTGGTGAAGTTTGGCCATGTTTCAAGTGGGGTAA 194  
 OY 1367 GTTATGTTGTTTGAAGAGATGTTGTACCAAGTAAACAACCTTATTCGCTTTTACT 1426  
 |||||||  
 DB 193 GTTATGTTGTTTGAAGAGATGTTGTACCAAGTAAACAACCTTATTCGCTTTTACT 134  
 OY 1427 TCTTGCTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
 |||||||  
 DB 133 TCTTGCTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 103

RESULT 13  
 AM774700 733 bp mRNA linear EST 07-SEP-2000  
 LOCUS EST333851 KV3 Medicago truncatula cDNA clone PKV3-23121, mRNA  
 DEFINITION sequence.  
 ACCESSION AM774700  
 VERSION AM774700.1 GI:7718617  
 KEYWORDS EST.  
 SOURCE Medicago truncatula  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;  
 Medicago.  
 1 (bases 1 to 733)  
 TITLE Vandenbosch, K., Hurl, J., Moore, J., Beremand, P., Peng, H., Ellis, L.,  
 Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, J.E. and  
 Fraser, C.M.  
 ESTs from roots of Medicago truncatula after Rhizobium inoculation  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Vandenbosch, K  
 Department of Biology  
 Texas A&M University  
 College Station, TX 77843-3258, USA  
 Tel: 409 845 7707  
 Fax: 409 845 2891  
 Email: Kate@mail.bio.tamu.edu



D	b	121	CAAGGATGTTCACGCCGGACAATAAGTGTCTCGAAGCAGAGAATTAAACGAGTAATGTT	180
O	y	352	TTTTGATGGCAAGCTCTCTGATATGTGAAGAACAGAACTTGCTAAAAGATACTCAAAAAAGA	411
D	b	181	TTTGATGGCAAGCTCTCTGATATGTGAAGAACAGAACTTGCTAAAAGATACTCAAAAAAGA	240
O	y	412	GATGATGCACCAAGATGTGCTAGCAGCTAGAGGTAGAGATTAAGATGGCATTTCAA	471
D	b	241	AATGAGCAACAGAGAGCTGACGAGAGCAAGCACTAGAGAGATTAAGATGGCATTTCAA	300
O	y	472	AAATGAGCAACAGAGCTGATTAAGGTCCACAGGCAACAAACAGATTTGAAGACACTA	531
D	b	301	AAATGAGCAACAGAGCTGATTAAGGTCCACAGGCAACAAACAGATTTGAAGACACTA	360
O	y	532	TTAAGCTTAATGGGGTTCCTCTGTATAGAGCACCTTCTGTAAGCAGAGCAAGATTTGCA	591
D	b	361	CTAAGACTGATGGGTTCCTCTGTGTGAGAGGCTCTGTGTGAGCAANATATCAATGTGCT	420
O	y	592	GCCCTTGGCATTAACATTAAGTGTGCTGCTGTGCTTCCAGAGATTAAGACCTCCCTTACT	651
D	b	421	GCCTTTGGCAAGATTCACAAAGGTGATTCOTGTTCGATCCGAAATATGACCTACTTACT	480
O	y	652	TTTGGGCTCCACGCTTCTCTGCTATTTATGATGCCAAGTTCCAGAGAAATACCTGT	710
D	b	481	TTTGGAGCTACACGCTTGTGCTGCTATTTGATGATCCCAAGTCCAGAGAAATACCTGT	539
<b>RESULT 15</b>				
AM288784	LOCUS	AM288784	414 bp mRNA linear EST J6-JUN-2000	
DEFINITION		707010C02.x5 707 - Mixed adult tissues from Walbot lab (Sk) Zea		
ACCESSION		mays cDNA, mRNA sequence.		
VERSION		AM288784		
KEYWORDS		EST .		
SOURCE		Zea mays.		
ORGANISM		Zea mays.		
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC		
		clade; Panicoideae; Andropogoneae; Zea.		
		1 (bases 1 to 414)		
REFERENCE		Maise ESTs from various cDNA libraries sequenced at Stanford		
AUTHORS		Walbot,V.		
TITLE		University		
JOURNAL		Unpublished (1999)		
COMMENT		Contact: Walbot V		
		Department of Biological Sciences		
		Stanford University		
		855 California Ave, Palo Alto, CA 94304, USA		
		Tel: 650 723 2227		
		Fax: 650 725 8221		
		Email: walbot@stanford.edu		
		Plate: 707010 row: C column: 02.		
FEATURES		Location/Qualifiers		
source		1..414		
		/organism="Zea mays"		
		/cultivar="W23"		
		/db_xref="taxon:4577"		
		/clone_lib="707 - Mixed adult tissues from Walbot lab (Sk)"		
		/tissue_type="cassel, kernel, silk, husk, root, leaf"		
		/dev_stage="adult"		
		/lab_host="DH10B"		
		/note="Organ: tassels, kernel, silk, husk, root, leaf; Vector: pBAD10, Site_1: EcoRI, cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4/2/1/1/1 (cassel, kernel, silk, husk, root, leaf). Unidirectionally cloned."		
BASE COUNT		120 a 82 c 101 g 109 t 2 others		
ORIGIN				
Query Match		27.0%; Score 399; DB 9; Length 414;		
Best Local Similarity		98.3%; Pzed No. 1e-57;		

	Matches	402,	Conservative	0;	Mismatches	7,	Indels	0;	Gaps	0;
QY	661	CCACGGTCCCTTCGTCATTTAAATGATCCCAAGTTCACAGAAAATACCTGATGACAAATTT	720							
Db	6	CCACGGTCCCTTCGTCATTTAAATGATCCCAAGTTCACAGAAAATACCTGATGAGAAATTT	65							
QY	721	GATGTTGCCAAGGTTTGGAGAGGCTTGACACACATGACGACCGATTCATGATTTGTGC	780							
Db	66	GATGTTGCCAAGGTTTGGAGAGGCTTGACACTCACATGAGACACAGNCTATTGATTTGTGC	125							
QY	781	ATCCCTGTTGGATGTGACTATTGTGATAGCATCAAGGATACGGGGGGCAACAGCTCTG	840							
Db	126	ATCCCTGTTGGATGTGACTATTGTGATAGCATCAAGGATATCGGGGGGCAACAGCTCTG	185							
QY	841	AAACTTATTCGTAAACATGGGTCCATACAGAACCATCTTGGAGATCTTAATAAAGACAG	900							
Db	186	AAACTTATTCGTAAACATGGGTCCATACAGAACCATCTTGGAGATCTTAATAAAGACAG	245							
QY	901	TATCAATTCCTGAGACTGCGCTTACACAGAAAGCTCGACCCCTTGTTCAGAGAGCCTAAT	960							
Db	246	TACCAATTCCTGAGACTGCGCTTACCAAGAAAGCTCGACCCCTTGNCTCAAGAGAGCCTAAT	305							
QY	961	GTCACTTTGGATATTCCTAGAGCTAAAATCGACTGCACCTGATGAGAGGCTTCATTAAGT	1020							
Db	306	GTCACATTTGGATATTCCTAGAGCTAAAATTCGACTGCACCTGATGAGAGGCTTCATTAAGT	365							
QY	1021	TTCCTGGTAAAGATATGATGTTTCATGAGAAATCGGAGACAAAGGCCA	1069							
Db	366	TTCCTGGTAAAGATATGATGTTTCACAGAAATTCGGCTGACAAAGGCCA	414							

Search completed: November 5, 2002, 13:37:11  
Job time : 1277.61 secs

QY 1021 TTCCTGGTAAAGATATGCTTTCATGAAGATCGGGTGACCAAGGCCA 1069  
|||||  
Db 366 TTCCTGGTAAAGATATGCTTTCACGAGAATCGGGTGACCAAGGCCA 414

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 11:03:23 : Search time 1855.32 Seconds  
(Without alignments)  
16501.458 Million cell updates/sec

Title: US-09-805-311-1  
Perfect score: 1463  
Sequence: 1 cacgagatactcgcgcgc.....aaaaaaaaaaaaaaaa 1463

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBank1 :  
1: gb\_ba :  
2: gb\_htg :  
3: gb\_in :  
4: gb\_cm :  
5: gb\_ov :  
6: gb\_pat :  
7: gb\_ph :  
8: gb\_pl :  
9: gb\_pr :  
10: gb\_ro :  
11: gb\_sts :  
12: gb\_sy :  
13: gb\_un :  
14: gb\_vl :  
15: em\_da :  
16: em\_fun :  
17: em\_hum :  
18: em\_in :  
19: em\_mu :  
20: em\_om :  
21: em\_or :  
22: em\_ov :  
23: em\_pat :  
24: em\_ph :  
25: em\_pl :  
26: em\_ro :  
27: em\_sts :  
28: em\_un :  
29: em\_vl :  
30: em\_htg\_hum :  
31: em\_htg\_inv :  
32: em\_htg\_other :  
33: em\_htgo\_inv :

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description
------------	-------------	-------	--------	----	----	-------------

1	1463	100.0	1463	6	AR152403	AR152403 Sequence
2	1360	93.0	1541	6	AR152404	AR152404 Sequence
3	1267	86.6	1381	6	AR152405	AR152405 Sequence
4	894	61.1	1478	6	AR152406	AR152406 Sequence
5	58	4.0	1354	8	AB021666	AB021666 Oryza sat
6	58	4.0	115907	2	AC104713	AC104713 Oryza sat
7	43	2.9	105137	2	HSJ104017	HSJ104017 Human DNA
8	40	2.7	2871	9	BC022967	BC022967 Homo sapi
9	40	2.7	89181	2	AC084072	AC084072 Homo sapi
10	39	2.7	288	6	AX185400	AX185400 Sequence
11	39	2.7	305	6	AX186826	AX186826 Sequence
12	39	2.7	224448	2	PFMAL4P4	PFMAL4P4 Plasmid
13	38	2.6	659	2	BC009510	BC009510 Homo sapi
14	38	2.6	691	6	AX011676	AX011676 Sequence
15	38	2.6	2228	4	HSM802710	HSM802710 Homo sapi
16	38	2.6	2240	4	AB042653	AB042653 Bos tauri
17	38	2.6	2291	9	BC001537	BC001537 Homo sapi
18	38	2.6	3315	3	AY069109	AY069109 Drosophila
19	38	2.6	67632	2	AC105102	AC105102 Homo sapi
20	38	2.6	38734	2	PFMAL1P2	PFMAL1P2 Plasmid
21	38	2.6	119077	9	AL354982	AL354982 Human DNA
22	38	2.6	134963	2	HS199L16	HS199L16 Human DNA
23	38	2.6	145069	2	AC012032	AC012032 Homo sapi
24	38	2.6	165894	9	AC092054	AC092054 Homo sapi
25	38	2.6	196560	2	AC027646	AC027646 Mus muscu
26	38	2.6	203250	9	AC020910	AC020910 Homo sapi
27	38	2.6	238728	2	AC025528	AC025528 Mus muscu
28	37	2.5	396	6	AX093307	AX093307 Sequence
29	37	2.5	445	9	BC009480	BC009480 Homo sapi
30	37	2.5	508	6	E01762	E01762 Variant aeg
31	37	2.5	508	6	E01765	E01765 DNA sequenc
32	37	2.5	826	6	E08516	E08516 DNA encodin
33	37	2.5	826	6	E12586	E12586 cDNA encodi
34	37	2.5	826	8	S65048	S65048 self-incomp
35	37	2.5	835	8	AF064030	AF064030 Helianthu
36	37	2.5	835	8	ATP27A	Y11792 A. thaliana
37	37	2.5	848	10 <th>BC018351</th> <th>BC018351 Mus muscu</th>	BC018351	BC018351 Mus muscu
38	37	2.5	925	3 <th>AEVA0440X</th> <th>L29571 Aequorea vi</th>	AEVA0440X	L29571 Aequorea vi
39	37	2.5	958	6 <th>AR108851</th> <th>AR108851 Sequence</th>	AR108851	AR108851 Sequence
40	37	2.5	958	6 <th>AR119320</th> <th>AR119320 Sequence</th>	AR119320	AR119320 Sequence
41	37	2.5	958	6 <th>AR151721</th> <th>AR151721 Sequence</th>	AR151721	AR151721 Sequence
42	37	2.5	958	6 <th>AX250563</th> <th>AX250563 Sequence</th>	AX250563	AX250563 Sequence
43	37	2.5	958	6 <th>E01573</th> <th>E01573 cDNA sequen</th>	E01573	E01573 cDNA sequen
44	37	2.5	958	6 <th>E01575</th> <th>E01575 cDNA sequen</th>	E01575	E01575 cDNA sequen
45	37	2.5	958	6 <th>E02319</th> <th>E02319 DNA encodin</th>	E02319	E02319 DNA encodin

# ALIGNMENTS

RESULT 1  
LOCUS AR152403 1463 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 1 from patent US 6232527.  
ACCESSION AR152403  
VERSION AR152403.1 GI:15118453  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1463)  
AUTHORS Manajan, P.B.  
TITLE Maize Rad2/FEN-1 orthologues and uses thereof  
JOURNAL Patent: US 6232527-A 1 15-MAY-2001;  
FEATURES  
source location/Qualifiers  
1..1463  
BASE COUNT 466 a 292 c 361 g 344 t  
ORIGIN  
Query Match 100.0%; Score 1463; DB 6; Length 1463;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGAGATATGCTCGGGCGCGGTTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCCA 60  
1 CAGAGATATGCTCGGGCGCGGTTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCCA 60  
Db 1 CAGAGATATGCTCGGGCGCGGTTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCCA 60  
QY 61 CCGGCCACAGCCGCGGCGGCGGTTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCCA 120  
61 CCGGCCACAGCCGCGGCGGCGGTTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCCA 120  
Db 61 CCGGCCACAGCCGCGGCGGCGGTTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCCA 120  
QY 121 AATGCCCAAGGCGGTTGAAGGAGCAAGTTCCAGAGCTCTTGGCCCGCAAAATCGC 180  
121 AATGCCCAAGGCGGTTGAAGGAGCAAGTTCCAGAGCTCTTGGCCCGCAAAATCGC 180  
Db 121 AATGCCCAAGGCGGTTGAAGGAGCAAGTTCCAGAGCTCTTGGCCCGCAAAATCGC 180  
QY 181 GTGAGCGCAGCATGAGCATATACCAATGTTCCGANTGTGATGTAAGAGCAGCATGAA 240  
181 GTGAGCGCAGCATGAGCATATACCAATGTTCCGANTGTGATGTAAGAGCAGCATGAA 240  
Db 181 GTGAGCGCAGCATGAGCATATACCAATGTTCCGANTGTGATGTAAGAGCAGCATGAA 240  
QY 241 ACTCTACAAATGAGTCTGTAAGTCACTATTTGCAAGCAATGTTCAACCGGACA 300  
241 ACTCTACAAATGAGTCTGTAAGTCACTATTTGCAAGCAATGTTCAACCGGACA 300  
Db 241 ACTCTACAAATGAGTCTGTAAGTCACTATTTGCAAGCAATGTTCAACCGGACA 300  
QY 301 ATAGAATTAAGTGAAGCGGCAATCAAGCAATTTATTTGTTGCAAGCCTCTGAT 360  
301 ATAGAATTAAGTGAAGCGGCAATCAAGCAATTTATTTGTTGCAAGCCTCTGAT 360  
Db 301 ATAGAATTAAGTGAAGCGGCAATCAAGCAATTTATTTGTTGCAAGCCTCTGAT 360  
QY 361 ATGAAGAACAGAGCTTCTTAAAGATCTCAAAAAGATGATGCAACCAAGATCTG 420  
361 ATGAAGAACAGAGCTTCTTAAAGATCTCAAAAAGATGATGCAACCAAGATCTG 420  
Db 361 ATGAAGAACAGAGCTTCTTAAAGATCTCAAAAAGATGATGCAACCAAGATCTG 420  
QY 421 ACTGAGGCAATGAGTGAAGATTAAGATGCGATTGAAAAATGAGCAAGAGACTGA 480  
421 ACTGAGGCAATGAGTGAAGATTAAGATGCGATTGAAAAATGAGCAAGAGACTGA 480  
Db 421 ACTGAGGCAATGAGTGAAGATTAAGATGCGATTGAAAAATGAGCAAGAGACTGA 480  
QY 481 AAGGTCAAGAGCAAGCAAGCAAGATTTGTAACGCGTATTAAGATTAAGAGGTTCC 540  
481 AAGGTCAAGAGCAAGCAAGCAAGATTTGTAACGCGTATTAAGATTAAGAGGTTCC 540  
Db 481 AAGGTCAAGAGCAAGCAAGCAAGATTTGTAACGCGTATTAAGATTAAGAGGTTCC 540  
QY 541 GTTGTAGAGGACCTTCTGTAAGCAAGCAAGCAAGATGCGACCTTTGCAATTAACGATAG 600  
541 GTTGTAGAGGACCTTCTGTAAGCAAGCAAGCAAGATGCGACCTTTGCAATTAACGATAG 600  
Db 541 GTTGTAGAGGACCTTCTGTAAGCAAGCAAGCAAGATGCGACCTTTGCAATTAACGATAG 600  
QY 601 GTTGTAGAGGACCTTCTGTAAGCAAGCAAGCAAGATGCGACCTTTGCAATTAACGATAG 660  
601 GTTGTAGAGGACCTTCTGTAAGCAAGCAAGCAAGATGCGACCTTTGCAATTAACGATAG 660  
Db 601 GTTGTAGAGGACCTTCTGTAAGCAAGCAAGCAAGATGCGACCTTTGCAATTAACGATAG 660  
QY 661 CGTCAATTAAGGATCAAGATTCAGAAATACCTGTGATGAATTTGATGTTGCCAAG 720  
661 CGTCAATTAAGGATCAAGATTCAGAAATACCTGTGATGAATTTGATGTTGCCAAG 720  
Db 661 CGTCAATTAAGGATCAAGATTCAGAAATACCTGTGATGAATTTGATGTTGCCAAG 720  
QY 721 GTTGTAGAGGACCTTCTGTAAGCAAGCAAGCAAGATGCGACCTTTGCAATTAACGATAG 780  
721 GTTGTAGAGGACCTTCTGTAAGCAAGCAAGCAAGATGCGACCTTTGCAATTAACGATAG 780  
Db 721 GTTGTAGAGGACCTTCTGTAAGCAAGCAAGCAAGATGCGACCTTTGCAATTAACGATAG 780  
QY 781 TGTGACTATTTGATAGATCAAGATTCAGAAATACCTGTGATGAATTTGATGTTGCCAAG 840  
781 TGTGACTATTTGATAGATCAAGATTCAGAAATACCTGTGATGAATTTGATGTTGCCAAG 840  
Db 781 TGTGACTATTTGATAGATCAAGATTCAGAAATACCTGTGATGAATTTGATGTTGCCAAG 840  
QY 841 CAACATGAGGCTTCAAGAGCAATCTTGAAGATTTAAGCAATTAACCAATTTCC 900  
841 CAACATGAGGCTTCAAGAGCAATCTTGAAGATTTAAGCAATTAACCAATTTCC 900  
Db 841 CAACATGAGGCTTCAAGAGCAATCTTGAAGATTTAAGCAATTAACCAATTTCC 900  
QY 901 GAGGACTGAGCTTACCAAGAGCTGAGCGCTTGTCAAGAGCCTAATGTACATTGAT 960  
901 GAGGACTGAGCTTACCAAGAGCTGAGCGCTTGTCAAGAGCCTAATGTACATTGAT 960  
Db 901 GAGGACTGAGCTTACCAAGAGCTGAGCGCTTGTCAAGAGCCTAATGTACATTGAT 960  
QY 961 ATTCCGAGCTTAATGAGCTGACCTGATGAGAGGCTTCAATTAAGTCTGAGTAA 1020  
961 ATTCCGAGCTTAATGAGCTGACCTGATGAGAGGCTTCAATTAAGTCTGAGTAA 1020  
Db 961 ATTCCGAGCTTAATGAGCTGACCTGATGAGAGGCTTCAATTAAGTCTGAGTAA 1020  
QY 1021 GATTAATGTTTCAAGAGAGTGGTGAACAAGGCCATAGAGATCAATCTGCCAAG 1080  
1021 GATTAATGTTTCAAGAGAGTGGTGAACAAGGCCATAGAGATCAATCTGCCAAG 1080  
Db 1021 GATTAATGTTTCAAGAGAGTGGTGAACAAGGCCATAGAGATCAATCTGCCAAG 1080

QY 1081 AATTAATGTCGCAAGAGAGTGAAGTCTTTTCAAGCCACTGCGCACCATGACGA 1140  
1081 AATTAATGTCGCAAGAGAGTGAAGTCTTTTCAAGCCACTGCGCACCATGACGA 1140  
Db 1081 AATTAATGTCGCAAGAGAGTGAAGTCTTTTCAAGCCACTGCGCACCATGACGA 1140  
QY 1141 CCGCTAAACGGAAGGAGACTTGGATTAACCAAGCAAGGAGCTGGCAACAGAAAC 1200  
1141 CCGCTAAACGGAAGGAGACTTGGATTAACCAAGCAAGGAGCTGGCAACAGAAAC 1200  
Db 1141 CCGCTAAACGGAAGGAGACTTGGATTAACCAAGCAAGGAGCTGGCAACAGAAAC 1200  
QY 1201 AAGGCTGTGGAAGAAATATCTGGATGCTGTGATGTAAGCAATGAGTACGAAG 1260  
1201 AAGGCTGTGGAAGAAATATCTGGATGCTGTGATGTAAGCAATGAGTACGAAG 1260  
Db 1201 AAGGCTGTGGAAGAAATATCTGGATGCTGTGATGTAAGCAATGAGTACGAAG 1260  
QY 1261 CAGGCTGTGGAAGAAATATCTGGATGCTGTGATGTAAGCAATGAGTACGAAG 1320  
1261 CAGGCTGTGGAAGAAATATCTGGATGCTGTGATGTAAGCAATGAGTACGAAG 1320  
Db 1261 CAGGCTGTGGAAGAAATATCTGGATGCTGTGATGTAAGCAATGAGTACGAAG 1320  
QY 1321 GGTAAAGTGTGCTCATGTTTCAAGCTGGGGTAAGTGTGTTGAAGAGATTGGG 1380  
1321 GGTAAAGTGTGCTCATGTTTCAAGCTGGGGTAAGTGTGTTGAAGAGATTGGG 1380  
Db 1321 GGTAAAGTGTGCTCATGTTTCAAGCTGGGGTAAGTGTGTTGAAGAGATTGGG 1380  
QY 1381 TACCAAGTAAACAAACTATGCTGTTTACTCTTGTCTTGAAGTAAAAA 1440  
1381 TACCAAGTAAACAAACTATGCTGTTTACTCTTGTCTTGAAGTAAAAA 1440  
Db 1381 TACCAAGTAAACAAACTATGCTGTTTACTCTTGTCTTGAAGTAAAAA 1440  
QY 1441 AAAAAAAAAAAAAAAAAA 1463  
1441 AAAAAAAAAAAAAAAAAA 1463  
Db 1441 AAAAAAAAAAAAAAAAAA 1463

RESULT 2  
AR152404 1541 bp DNA linear PAT 08-AUG-2001  
LOCUS AR152404  
DEFINITION Sequence 3 from patent US 6232527.  
ACCESSION AR152404  
VERSION AR152404.1 GI:15118454  
KEYWORDS  
Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1541)  
AUTHORS Mahajan, P. B.  
TITLE Maize Red2/FEN-1 orthologues and uses thereof  
JOURNAL Patent: US 6232527-A 3 15-MAY-2001;  
FEATURES  
Source  
1. AR1541  
/organism="unknown"  
BASE COUNT 473 a 308 c 377 g 383 t  
ORIGIN  
Query Match 93.0%; Score 1360; DB 6; Length 1541;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 GCGGTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCACCGCCAGCGCCGAGA 80  
21 GCGGTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCACCGCCAGCGCCGAGA 80  
Db 21 GCGGTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCACCGCCAGCGCCGAGA 80  
QY 15 GCGGTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCACCGCCAGCGCCGAGA 74  
15 GCGGTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCACCGCCAGCGCCGAGA 74  
Db 15 GCGGTTCTTGCGCCACTCCGGCTACAGCCGCCCGCCACCGCCAGCGCCGAGA 74  
QY 81 CGAGATGGGATCAAGGAGTTTGACGAACCTGCTGGGAGCAATGCGCCAGGCGATGA 140  
81 CGAGATGGGATCAAGGAGTTTGACGAACCTGCTGGGAGCAATGCGCCAGGCGATGA 140  
Db 81 CGAGATGGGATCAAGGAGTTTGACGAACCTGCTGGGAGCAATGCGCCAGGCGATGA 140  
QY 75 CGAGATGGGATCAAGGAGTTTGACGAACCTGCTGGGAGCAATGCGCCAGGCGATGA 134  
75 CGAGATGGGATCAAGGAGTTTGACGAACCTGCTGGGAGCAATGCGCCAGGCGATGA 134  
Db 75 CGAGATGGGATCAAGGAGTTTGACGAACCTGCTGGGAGCAATGCGCCAGGCGATGA 134  
QY 141 GAGCAGAAAGTTCGAGAGTCTGAGCGGCGCAAAATCGCGCTGAGCGCCAGCATGAGCAT 200  
141 GAGCAGAAAGTTCGAGAGTCTGAGCGGCGCAAAATCGCGCTGAGCGCCAGCATGAGCAT 200  
Db 141 GAGCAGAAAGTTCGAGAGTCTGAGCGGCGCAAAATCGCGCTGAGCGCCAGCATGAGCAT 200  
QY 201 ATTCAGATTCGATGTTGATGTAAGGAGCAAGGCAATGCAATTAAGTACGATG 260  
201 ATTCAGATTCGATGTTGATGTAAGGAGCAAGGCAATGCAATTAAGTACGATG 260  
Db 201 ATTCAGATTCGATGTTGATGTAAGGAGCAAGGCAATGCAATTAAGTACGATG 260  
QY 195 ATTCAGATTCGATGTTGATGTAAGGAGCAAGGCAATGCAATTAAGTACGATG 254  
195 ATTCAGATTCGATGTTGATGTAAGGAGCAAGGCAATGCAATTAAGTACGATG 254  
Db 195 ATTCAGATTCGATGTTGATGTAAGGAGCAAGGCAATGCAATTAAGTACGATG 254  
QY 261 TGAAGTCACTAGTCAATTTCAAGAGATGTTCAACCGCAATTAAGTACTGAGAGCGG 320  
261 TGAAGTCACTAGTCAATTTCAAGAGATGTTCAACCGCAATTAAGTACTGAGAGCGG 320  
Db 261 TGAAGTCACTAGTCAATTTCAAGAGATGTTCAACCGCAATTAAGTACTGAGAGCGG 320

```

QY 321 AATCAAGCAGTTATGTTTGTGTCAGACCTCCTGATATGAAGAAACAAGCTTGC 380
    |||
Db 315 AATCAAGCAGTTATGTTTGTGTCAGACCTCCTGATATGAAGAAACAAGCTTGC 374
QY 381 TAAAGATCTCAAAAGAGATGATGACCAAGAATCTGACTGAGCGAGTAGAGTGG 440
    |||
Db 375 TAAAGATCTCAAAAGAGATGATGACCAAGAATCTGACTGAGCGAGTAGAGTGG 434
QY 441 AGATAAAGATGCGATTGAAAAATTGACAGAGAGAGCTGTAAGGTCACAAAGCAACAA 500
    |||
Db 435 AGATAAAGATGCGATTGAAAAATTGACAGAGAGAGCTGTAAGGTCACAAAGCAACAA 494
QY 501 CGAAGATTTGAACGGCTATTGAACATTATGAGGGTCTCTGTGTGAAGCAGCTTCTGA 560
    |||
Db 495 CGAAGATTTGAACGGCTATTGAACATTATGAGGGTCTCTGTGTGAAGCAGCTTCTGA 554
QY 561 AGCAGAGCAGAAATGTCAGCCCTTTCATTAACGATAGAGTGTTCCTGTGCTTGA 620
    |||
Db 555 AGCAGAGCAGAAATGTCAGCCCTTTCATTAACGATAGAGTGTTCCTGTGCTTGA 614
QY 621 AGATATGACTCCCTTACTTTTGGGGCTCCAGGGTTCCTTCGTATTATGATCCAG 680
    |||
Db 615 AGATATGACTCCCTTACTTTTGGGGCTCCAGGGTTCCTTCGTATTATGATCCAG 674
QY 681 TTCCAAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGTTTGGAGAGCTTGA 740
    |||
Db 675 TTCCAAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGTTTGGAGAGCTTGA 734
QY 741 CACCATGAGCAGTTCATTAATTTGTCATCTGTGATGATGATGATGATGATGATGAT 800
    |||
Db 735 CACCATGAGCAGTTCATTAATTTGTCATCTGTGATGATGATGATGATGATGATGAT 794
QY 801 CAAAGGATATGGGGGGCAAAACAGCTTGAAACTTATTCGTCACATGGGTCATAGAAG 860
    |||
Db 795 CAAAGGATATGGGGGGCAAAACAGCTTGAAACTTATTCGTCACATGGGTCATAGAAG 854
QY 861 CATCTTGAGAACTCTTAATAAGACATATCAATTCCTGAGACCTGCGCTTACCAAGA 920
    |||
Db 855 CATCTTGAGAACTCTTAATAAGACATATCAATTCCTGAGACCTGCGCTTACCAAGA 914
QY 921 AGCTGACGCTTGTCAAGAGGCTTAATGTCATGATGATATTCCTGAGCTTAATAAGT 980
    |||
Db 915 AGCTGACGCTTGTCAAGAGGCTTAATGTCATGATGATATTCCTGAGCTTAATAAGT 974
QY 981 TGCACCTGATGAGAGGGTCTCATAGTTCTGTGTGTAAGATATGTTTCAACGAAGA 1040
    |||
Db 975 TGCACCTGATGAGAGGGTCTCATAGTTCTGTGTGTAAGATATGTTTCAACGAAGA 1034
QY 1041 TCGGGTGACAAAGGCCATAGAAAGATCAATTCGCCAAGAATTAATCGTCGAAGGAAG 1100
    |||
Db 1035 TCGGGTGACAAAGGCCATAGAAAGATCAATTCGCCAAGAATTAATCGTCGAAGGAAG 1094
QY 1101 ACTGAGCTCTTTTCAAGCCAACTGCCACACATGACACCGCTTAATAAGGAGAGAC 1160
    |||
Db 1095 ACTGAGCTCTTTTCAAGCCAACTGCCACACATGACACCGCTTAATAAGGAGAGAC 1154
QY 1161 TTGGGATTAACAAAGCAAGGCACTGTGGAACAAAGAAACAAGGCTGTGGAAGGAAGA 1220
    |||
Db 1155 TTGGGATTAACAAAGCAAGGCACTGTGGAACAAAGAAACAAGGCTGTGGAAGGAAGA 1214
QY 1221 ATAATTTGATGCTTGTATGATTAACATACAGACTACGAAAGCAGCGGGGTGATCACTT 1280
    |||
Db 1215 ATAATTTGATGCTTGTATGATTAACATACAGACTACGAAAGCAGCGGGGTGATCACTT 1274
QY 1281 CGCTTGATTAATTAACCTCCCTGTTTAACTAGAGACTTTGTAAGATTTGCTCATGTT 1340
    |||
Db 1275 CGCTTGATTAATTAACCTCCCTGTTTAACTAGAGACTTTGTAAGATTTGCTCATGTT 1334
QY 1341 TCAAGCTGGGGTAAAGTAAAGTGTGTTGAAGAGATGGTGTACCAAGTAACAAACTTAT 1400
    |||
Db 1335 TCAAGCTGGGGTAAAGTAAAGTGTGTTGAAGAGATGGTGTACCAAGTAACAAACTTAT 1394
QY 1401 CGCTGTTTTTACTCTTGCTCTTGAAGTA 1431

```

```

Db 1395 CGCTGTTTTTACTCTTGCTCTTGAAGTA 1425
    |||
RESULT 3
LOCUS AR152405 AR152405 1381 bp DNA Linear PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6232527.
ACCESSION AR152405
VERSION AR152405.1 GI:15118455
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1381)
AUTHORS Mahajan,P.B.
TITLE Maize Rad2/FEEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 5 15-MAY-2001;
FEATURES
source 1. 1381
BASE COUNT 441 a 269 c 346 g 325 t
ORIGIN
Query Match 86.6%; Score 1267; DB 6; Length 1381;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 64 GCCACAGCCGCCGAGACAGATGGGATCAAGGTTTGAAGAACTGTGGCGACAT 123
    |||
Db 16 GCCACAGCCGCCGAGACAGATGGGATCAAGGTTTGAAGAACTGTGGCGACAT 75
QY 124 GCGCCCAAGGGGATGAAGAGACAGAGTTCGAGAGCTACTCGGCCGAAATGCGCGTC 183
    |||
Db 76 GCGCCCAAGGGGATGAAGAGACAGAGTTCGAGAGCTACTCGGCCGAAATGCGCGTC 135
QY 184 GAGCCACGATGACATATACCATTCCTGATGTTAGTTGGAAGAGACAGGATGAAACT 243
    |||
Db 136 GAGCCACGATGACATATACCATTCCTGATGTTAGTTGGAAGAGACAGGATGAAACT 195
QY 244 CTCACAAATGAAGGTGGTGAAGTCACTAGTCATTTGCAAGAAATGTTCAACCGGACATA 303
    |||
Db 196 CTCACAAATGAAGGTGGTGAAGTCACTAGTCATTTGCAAGAAATGTTCAACCGGACATA 255
QY 304 AGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTGTGAGCAAGCTCTGATATG 363
    |||
Db 256 AGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTGTGAGCAAGCTCTGATATG 315
QY 364 AAGAAACAAGAGCTTGTCTAAAGATCTCAAAAAGAGATGACCAACCAAGATCTGACT 423
    |||
Db 316 AAGAAACAAGAGCTTGTCTAAAGATCTCAAAAAGAGATGACCAACCAAGATCTGACT 375
QY 424 GAGCAGTAGAGGTAGAGATTAAGATGCAATTTGAAGAAATTTGACCAAGAGACTGTAAG 483
    |||
Db 376 GAGCAGTAGAGGTAGAGATTAAGATGCAATTTGAAGAAATTTGACCAAGAGACTGTAAG 435
QY 484 GTCAACAAGGCAACAACAAGAAATTTGTAACGGCTATTAAAGACTTATGAGGGTCTCTGT 543
    |||
Db 436 GTCAACAAGGCAACAACAAGAAATTTGTAACGGCTATTAAAGACTTATGAGGGTCTCTGT 495
QY 544 GTGAGAGCACTTCTGAAGCAAGAAATGTGACAGCCCTTTCGATTAACGATAGGTG 603
    |||
Db 496 GTGAGAGCACTTCTGAAGCAAGAAATGTGACAGCCCTTTCGATTAACGATAGGTG 555
QY 604 TTGCTGTGCTTCAAGAGATATGAGTCCCTTACTTTTGGGGCTCAGGGTCTCTGCT 663
    |||
Db 556 TTGCTGTGCTTCAAGAGATATGAGTCCCTTACTTTTGGGGCTCAGGGTCTCTGCT 615
QY 664 CATTTAATGATCAAGATTTCAAGAAATACCTGTATGGAATTTGATGTTGCAAGGTT 723
    |||
Db 616 CATTTAATGATCAAGATTTCAAGAAATACCTGTATGGAATTTGATGTTGCAAGGTT 675
QY 724 TTGGAGAGGCTTAACTCAACATGAGACAGTTCAATTGATTTGTGCATCTGTGTGATGT 783

```

|||||  
Db 676 TTGAGAGAGCTTAACCTACACATGAGACCGTTCATTGATTGGCGCTCGTGTGATGT 735  
QY 784 GACTATTTGATAGATCAAAAGATATCGGGGGGCAACAGCTCTGAATCTTATTCGTCAA 843  
Db 736 GACTATTTGATAGATCAAAAGATATCGGGGGGCAACAGCTCTGAATCTTATTCGTCAA 795  
QY 844 CATGGGTCCATAGAAAGCATCTTGGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 903  
Db 796 CATGGGTCCATAGAAAGCATCTTGGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 855  
QY 904 GACGGGCTTACCAAGAGCTGACGCTTGTTCAGAGGCTTATGCAATGATATT 963  
Db 856 GACTGGCTTACCAAGAGCTGACGCTTGTTCAGAGGCTTATGCAATGATATT 915  
QY 964 CCTGAGCTAAATGAGATGACCTGATGAGAGGGTCTCTAATAGTTTCTGTGTAAGAT 1023  
Db 916 CCTGAGCTAAATGAGATGACCTGATGAGAGGGTCTCTAATAGTTTCTGTGTAAGAT 975  
QY 1024 AATGGTTCAACGAAGATCGGGTGACAAAGGCGCATGAGAAATCAATCTGCCAAGAT 1083  
Db 976 AATGGTTCAACGAAGATCGGGTGACAAAGGCGCATGAGAAATCAATCTGCCAAGAT 1035  
QY 1084 AAATCGTCGCAAGAGAGCTCGAGTCTTTTTCAGCCAACTGCCACCATGACGACCG 1143  
Db 1036 AAATCGTCGCAAGAGAGCTCGAGTCTTTTTCAGCCAACTGCCACCATGACGACCG 1095  
QY 1144 CTAAACGGAAGAGAGCTCGGATTAACAGCAAGGAGCTGCGAACAAGAAACAAAG 1203  
Db 1096 CTAAACGGAAGAGAGCTCGGATTAACAGCAAGGAGCTGCGAACAAGAAACAAAG 1155  
QY 1204 GCTGGTGAAGAAGAATATCTGTGATGCTTGTGATGATACACTAGACAGCAAGAGAG 1263  
Db 1156 GCTGGTGAAGAAGAATATCTGTGATGCTTGTGATGATACACTAGACAGCAAGAGAG 1215  
QY 1264 CGGTGGGATGATCACTTCGTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTTGGT 1323  
Db 1216 CGGTGGGATGATCACTTCGTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTTGGT 1275  
QY 1324 AAAAGTT 1330  
Db 1276 AAAAGTT 1282

RESULT 4  
ARI52406 1478 bp DNA linear PAT 08-AUG-2001  
LOCUS ARI52406  
DEFINITION Sequence 7 from patent US 6232527.  
ACCESSION ARI52406  
VERSION ARI52406.1 GI:15118456  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1478)  
AUTHORS Mahajan, P.B.  
TITLE Maize Rad2/REN-1 orthologues and uses thereof  
JOURNAL Patent: US 6232527-A 7 15-MAY-2001.  
FEATURES  
source location/Qualifiers  
1..1478  
BASE COUNT 463 a 302 c 365 g 348 t  
ORIGIN

Query Match 61.1%; Score 894; DB 6; Length 1478;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1244; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 20 CGCGGTTTCTTGGCAGCACTCGGCTCAGCCGCGCCGCGACCGCAGCCGCGCAG 79  
Db 32 CGCGGTTTCTTGGCAGCACTCGGCTCAGCCGCGCCGCGACCGCAGCCGCGCAG 91  
QY 80 ACGAGATGGGATCAAGGGTTTGACGAACCTGTCGCGACATGCGCCCAAGGCGATGA 139

|||||  
Db 92 ACGAGATGGGATCAAGGTTTGACGAACCTGCTGGCGGACCAATGCCCCAAAGGCGATGA 151  
QY 140 AGGAGCAGAAAGTTGAGAGAGTACTTGGCGGCAAAATCGCGTGCAGCGCAGATGACA 199  
Db 152 AGGAGCAGAAAGTTGAGAGAGTACTTGGCGGCAAAATCGCGTGCAGCGCAGATGACA 211  
QY 200 TATACAGTTCTGATTTGATTTGGAAAGCAGCGCATGGAACCTCAGCAAAATGAACCTG 259  
Db 212 TCTACAGTTCTGATTTGATTTGGAAAGCAGCGCATGGAACCTCAGCAAAATGAACCTG 271  
QY 260 GTGAAGTACTAGTCAATTTGCAAGAAATGTTCAACCGGACAAATAGATTACTGGAAACGG 319  
Db 272 GTGAAGTACTAGTCAATTTGCAAGAAATGTTCAACCGGACAAATAGATTACTGGAAACGG 331  
QY 320 GAATCAAGCCAGTTTATGTTTATGATGGCAAGCCTCTGATATGAGAAACAGAGCTTG 379  
Db 332 GAATCAAGCCAGTTTATGTTTATGATGGCAAGCCTCTGATATGAGAAACAGAGCTTG 391  
QY 380 CTAAAGATTAAGTCAAAAGATGATGCAACCAAGATCTGAGTGAAGTGAAGCTAG 439  
Db 392 CTAAAGATTAAGTCAAAAGATGATGCAACCAAGATCTGAGTGAAGTGAAGCTAG 451  
QY 440 GAGATTAAGATGCGATTGAAAAATTTAGCAAGAGAGCTGTAAAGTCAACAGGCAACACA 489  
Db 452 GAGATTAAGATGCGATTGAAAAATTTAGCAAGAGAGCTGTAAAGTCAACAGGCAACACA 511  
QY 500 ACGAAGATTTGAACGGCTATTAAGACTTATGGGGTTCCCTGTTGTAGAGGCACTCTTG 559  
Db 512 ACGAAGATTTGAACGGCTATTAAGACTTATGGGGTTCCCTGTTGTAGAGGCACTCTTG 571  
QY 560 AAGCAAGCAAGAAATGTCACAGCCTTTGCAATTAAGCAATGATGATGCTGTTGCTTACG 619  
Db 572 AAGCAAGCAAGAAATGTCACAGCCTTTGCAATTAAGCAATGATGATGCTGTTGCTTACG 631  
QY 620 AAGATATGAGTCCCTTACTTTTGGGGCTCCAGGTTCTTCTGCAATTTAATGATCAA 679  
Db 632 AAGATATGAGTCCCTTACTTTTGGGGCTCCAGGTTCTTCTGCAATTTAATGATCAA 691  
QY 680 GTTCCAGAAATACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729  
Db 692 GTTCCAGAAATACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 751  
QY 740 TCACCATGGACAGTTCAATGATTTGTGATCCTGTGTGATGATGATGATGATGATGATGATGAT 799  
Db 752 TCACCATGGACAGTTCAATGATTTGTGATCCTGTGTGATGATGATGATGATGATGATGATGAT 811  
QY 800 TCAAAAGTATCGGGGGGCAACAGCTCTGAAACTTATTCGACACATGGGTCCATAGAAA 859  
Db 812 TCAAAAGTATCGGGGGGCAACAGCTCTGAAACTTATTCGACACATGGGTCCATAGAAA 871  
QY 860 GCATCTTGAGATCTTATTAAGACAGATATCAAAATTCGAGAGCTGAGCTTACCAG 919  
Db 872 GCATCTTGAGATCTTATTAAGACAGATATCAAAATTCGAGAGCTGAGCTTACCAG 931  
QY 920 AAGCTGAGAGCTTGTTCAGAGACCTTAATGTCACATTTGATGATGATGATGATGATGATGAT 979  
Db 932 AAGCTGAGAGCTTGTTCAGAGACCTTAATGTCACATTTGATGATGATGATGATGATGATGATGAT 991  
QY 980 CTGCACTGATGAGAGGGGCTCATTAAGTTTCTGCTGATTAAGATTAAGGCTTCAACGAG 1039  
Db 992 CTGCACTGATGAGAGGGGCTCATTAAGTTTCTGCTGATTAAGATTAAGGCTTCAATGAG 1051  
QY 1040 ATCGGGTGCAAAAGGCGCATAGAGATCAATCGCCCAAGATTAATCGTCGCAAGGAA 1099  
Db 1052 ATCGGGTGCAAAAGGCGCATAGAGATCAATCGCCCAAGATTAATCGTCGCAAGGAA 1111  
QY 1100 GACTCGAGTCTTTTTCAGGCCACTGCGACCATTCGACCGCTTAACGAGAGAGA 1159  
Db 1112 GACTCGAGTCTTTTTCAGGCCACTGCGACCATTCGACCGCTTAACGAGAGAGA 1171  
QY 1160 CTTGGATTAACCAAGCAGAGCTGCGAACCAAGAAAGGCTGCTGGAAGAGA 1219



db	1172	CTTCGGATTAACCAAGCAAGCAGAGTGGTGAACAAGAAACCAAGGCGTGGTGAAGAAGA	1231
Oy	1220	AATAATCTTGATGCTGTGATGTACACACTACACACTACGAAGCAGCGGTGGC	1270
D	1232	AATTAATCTTGATGCTGTGATGTACACACTACACACTACGAAGCAGCGGTGGC	1282
RESULT 5			
AB021666			
LOCUS			
DEFINITION	AB021666	1354 bp mRNA linear	PLN 25-MAR-2000
ACCESSION	AB021666		
VERSION	AB021666.2	GI:4587224	
KEYWORDS	FEN-1; endonuclease.		
SOURCE	Oryza sativa cDNA to mRNA.		
ORGANISM	Oryza sativa		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE			
AUTHORS	Kimura,S., Ueda,T., Hatanaka,M., Takenouchi,M., Hashimoto,J. and Sakaguchi,K.		
TITLE	Plant homologue of flap endonuclease-1: molecular cloning, characterization, and evidence of expression in meristematic tissues		
JOURNAL	Plant Mol. Biol. 42 (3), 415-427 (2000)		
MEDLINE	20256470		
REFERENCE	2 (bases 1 to 1354)		
AUTHORS	Kimura,S., Hashimoto,J. and Sakaguchi,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-DEC-1998) Satsuke Kimura, Science University of Tokyo, Dept. of Applied Biological Science; 2641 Yamazaki, Noda, Chiba 278-8510, Japan (E-mail:j5498703@ed.noda.sut.ac.jp, Tel:81-471-24-1501(ex.3419), Fax:81-471-23-9767)		
COMMENT	On Apr 17, 1999 this sequence version replaced gi:4062866.		
FEATURES			
source	location/Qualifiers		
	1..1354		
gene	/organism="Oryza sativa"		
	/db_xref="taxon:4530"		
	67..1209		
CDS	/gene="OsFEN-1"		
	67..1209		
	/gene="OsFEN-1"		
	/note="endonuclease"		
	/codon_start=1		
	/product="FEN-1"		
	/protein_1g="AAA36171.1"		
	/db_xref="GI:4587225"		
	/translation="WGIKGLKLIADNPAMKEOKFESYFGRIAYDASISYFOLITVVRGMEITLTINEAGEVYTHLQGFENFKITLLEAGTKRPYVVDGKPPDLKQOELAKRRIKREKATKELTEAVAEEDKDAIEKFSKRVTKYKQKNEBCKRLLELVPAVEACEAEKAEALCINDMVAVASEDMDSITFEGAPFELHLNDPSSKKIPVNEEVAKVELEELTEIMDFIDILCISGDCYDSTKIGIGOTATLKLROHGSISTEILEININDRYOIPEDMWFYOEAEARKLEPNVTLIDIPBLKMANPDEGILVEFLVENGFNDRYTKAILEIKFEAKNNKSSQGLSEFKEPVYSTVPLKRDTSKRPKANANKTKSAGSKKK"		
	1354		
	/note="17 A nucleotides"		
BASE COUNT	415 a 286 c 331 g 322 t		
ORIGIN			
	Query Match	4.0%; Score 58; DB 8; Length 1354;	
	Best Local Similarity 100.0%; Pred. No. 1.5e-21;		
Matches	58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Oy	113	TGGCGGCAATGGCCCAAGCGATGAGAGAGACAGAACTTGGAGAGCTACTTGGCCGC	170
D	95	TGGCGGCAATGGCGCCCAAGCGATGAGAGAGACAGAACTTGGAGAGCTACTTGGCCGC	152
RESULT 6			
LOCUS	AC104713	115907 bp DNA linear	HTG 20-DEC-2001
DEFINITION	Oryza sativa chromosome 5 clone OUI362611, *** SEQUENCING IN		

```

ACCESSION      PROGRESS  ***, 3 ordered pieces.
VERSION        AC104713
KEYWORDS       HTG: HTGS_PHASE2.
SOURCE         Oryza sativa.
ORGANISM       Oryza sativa
REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS       Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
               Ebracteoidae; Oryzaceae; Oryza.
REFERENCE      1 (bases 1 to 115907)
AUTHORS       Chow,T.-Y., Hsing,Y.-I.C., Chen,C.-S., Chen,H.-H., Wu,H.-P.,
               Liu,S.-M., Chao,Y.-T., Chang,S.-J., Chen,T.-R., Chen,Y.-D.,
               Chow,M.-H.J., Hong,Y.-C., Hsiung,J.-N., Hsu,C.-H., Huang,J.-J.,
               Kau,P.-I., Lee,M.-C., Leu,H.-L., Lin,S.-J., Wu,L.-F., and Shaw,J.-F.
               Oryza sativa BAC OJ136ZG11 genomic sequence
               Unpublished
               2 (bases 1 to 115907)
               Chow,T.-Y. and Hsing,Y.-I.C.
TITLE          Direct Submission
JOURNAL        Submitted (20-DEC-2001) Institute of Botany, Academia Sinica, 128,
               Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
REFERENCE      The nucleotide sequence of this BAC clone was generated by
AUTHORS       Combining Monsanto and ASPGC-Taiwan sequencing data.
TITLE          * NOTE: This is a 'working draft' sequence. It currently
               * consists of 3 contigs. Gaps between the contigs
               * are represented as runs of N. The order of the pieces
               * is believed to be correct as given, however the sizes
               * of the gaps between them are based on estimates that have
               * provided by the submitter.
               * This sequence will be replaced
               * by the finished sequence as soon as it is available and
               * the accession number will be preserved.
               * 1
               * 3720 3739: contig of 3719 bp in length
               * 3720 3739: gap of unknown length
               * 3740 110847: contig of 107108 bp in length
               * 110848 110867: gap of unknown length
               * 110868 115907: contig of 5040 bp in length.
               * location/Qualifiers
               1.115907
               /organism="Oryza sativa"
               /cultivar="Nipponbare"
               /db_xref="taxon:4530"
               /chromosome="5"
               /clone="OJ136ZG11"
BASE COUNT    32362 a 24964 c 25806 g 32735 t 40 others
ORIGIN
Query Match 4.0%: Score 58; DB 2; Length 115907;
Best Local Similarity 100.0%; Pred. No. 1,8e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 113 TGGCGCAGCATGCGCCAGCGCATGACAGAGACACAAAGTTGGAGAGCTACTTGGCGCG 170
|||||
Db 8213 TGGCGGCAATGGCGCCAGCGCATGACAGAGACACAAAGTTGGAGAGCTACTTGGCGCG 8156

RESULT 7
HSJ104017/c 105137 bp DNA linear PRI 15-FEB-2000
LOCUS       Human DNA sequence from clone RPI-104017 on chromosome 6q16.1-16.3
DEFINITION  Contains STSs and GSSs, complete sequence.
ACCESSION   AL080316
VERSION     AL080316.8 GI:5650657
KEYWORDS    HMG.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 105137)
AUTHORS     Tracey,A.
TITLE       Direct Submission
JOURNAL     Submitted (14-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone

```

## COMMENT

requests: clonerequest@sanger.ac.uk  
On Jul 29, 1999 this sequence version replaced gi:5566542.  
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the feature key is ambiguous, there is an annotation using the 'unsure' feature key.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at  
<http://www.sanger.ac.uk/projects/Celegans/wormpep> RPI-104017 is from the library RPI-1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:pcypac2>

This sequence is the entire insert of clone RPI-104017. This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/MGP/Ch6>.

## FEATURES

```
source
location/Qualifiers
1..105137
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="q16.1-16.3"
/clone_lib="RPI-1"
/clone="RPI-104017"
/complement(372..811)
/note="match: GSS: Em:A0006718"
1037..1080
/note="22 copies 2 mer ta 88% conserved"
1153..1456
/note="AluSx repeat: matches 1..301 of consensus"
2365..2702
/note="L2 repeat: matches 1..2750 of consensus"
3579..3733
/note="L2 repeat: matches 2363..2750 of consensus"
4961..5272
/note="L2 repeat: matches 2041..2201 of consensus"
5545..5648
/note="AluY repeat: matches 1..311 of consensus"
7388..7668
/note="52 copies 2 mer ta 71% conserved"
7871..8066
/note="AluJo repeat: matches 2..283 of consensus"
/note="MER58A repeat: matches 10..224 of consensus"
/complement(10286..10857)
/note="match: GSS: Em:B02911"
11075..11346
/note="L1ME3 repeat: matches 5657..5940 of consensus"
/complement(13059..13460)
/note="match: STS: Em:G21604"
14002..14199
/note="MER58C repeat: matches 3..89 of consensus"
17080..17249
/note="L1MC1 repeat: matches 6165..6323 of consensus"
17425..17538
/note="MIR repeat: matches 131..245 of consensus"
/complement(17580..17784)
/note="match: STS: Em:G43022"
18332..18525
/note="AluJo repeat: matches 87..295 of consensus"
18919..19338
/note="L2 repeat: matches 1868..2326 of consensus"
19344..19733
```

```
/note="MIR1B repeat: matches 12..389 of consensus"
20175..20286
/note="56 copies 2 mer at 64% conserved"
20512..20976
/note="L2 repeat: matches 2264..2750 of consensus"
21113..21217
/note="L2 repeat: matches 1233..1341 of consensus"
21331..21517
/note="L1MC repeat: matches 1768..1953 of consensus"
21882..22020
/note="L1ME repeat: matches 419..560 of consensus"
22615..22710
/note="48 copies 2 mer tt 62% conserved"
23360..23417
/note="29 copies 2 mer tt 69% conserved"
24769..25327
/note="HERV16 repeat: matches 6..548 of consensus"
25328..26196
/note="MER6 repeat: matches 1..864 of consensus"
26197..26818
/note="HERV16 repeat: matches 548..1219 of consensus"
26964..27296
/note="L1R16A repeat: matches 95..450 of consensus"
27297..27643
/note="HERV1 repeat: matches 3587..3946 of consensus"
28215..28338
/note="MST1 repeat: matches 284..422 of consensus"
28589..29009
/note="match: GSS: Em:A0308663"
28617..28812
/note="MST1 repeat: matches 1..178 of consensus"
29010..29383
/note="MIR1B repeat: matches 1..388 of consensus"
30078..30297
/note="L1R16A repeat: matches 15..229 of consensus"
32881..33162
/note="L1MC5 repeat: matches 7649..7926 of consensus"
33433..33616
/note="MIR repeat: matches 40..260 of consensus"
/complement(33617..33834)
/note="match: GSS: Em:A0240281"
33899..34036
/note="MIR1 repeat: matches 269..404 of consensus"
37901..38084
/note="AluSg/X repeat: matches 83..264 of consensus"
38096..38119
/note="L2 copies 2 mer ac 95% conserved"
38855..38952
/note="MIR repeat: matches 85..184 of consensus"
39820..40101
/note="AluY repeat: matches 1..299 of consensus"
40284..41588
/note="L1P17 repeat: matches 2138..3447 of consensus"
41588..44294
/note="L1P17 repeat: matches 3441..6142 of consensus"
45641..46728
/note="L1ME1 repeat: matches 4994..6109 of consensus"
46737..47808
/note="TIGER1 repeat: matches 1350..2418 of consensus"
47815..48092
/note="Charlfield repeat: matches 70..386 of consensus"
48179..48335
/note="L1P1 repeat: matches 5999..6155 of consensus"
48338..48369
/note="L2 copies 2 mer aa 93% conserved"
48370..48882
/note="L1M4 repeat: matches 4272..4790 of consensus"
48927..49175
/note="L1MC repeat: matches 2146..2412 of consensus"
49180..49590
/note="L1M3 repeat: matches 7310..7739 of consensus"
49630..49717
/note="L1MC4 repeat: matches 7219..7313 of consensus"
```

```

repeat_region 49732..50676
/note="LMC repeat: matches 1514..2152 of consensus"
repeat_region 50716..50836
/note="LMC repeat: matches 1287..1413 of consensus"
repeat_region 50849..51871
/note="LMC1 repeat: matches 5287..6316 of consensus"
repeat_region 51874..52239
/note="LMC repeat: matches 929..1302 of consensus"
repeat_region 52268..52519
/note="LMC repeat: matches 523..780 of consensus"
repeat_region 53098..53206
/note="Alu repeat: matches 197..305 of consensus"
repeat_region 53208..53388
/note="MIR repeat: matches 41..231 of consensus"
repeat_region 53414..53830
/note="L1P8 repeat: matches 5746..6163 of consensus"
repeat_region 54168..54264
/note="MIR repeat: matches 2..111 of consensus"
misc_feature complement(54484..55028)
/note="match: GSS: Em:A0318586"
repeat_region 55956..56266
/note="AluX repeat: matches 1..312 of consensus"
repeat_region 56471..56682
/note="MIR repeat: matches 6..232 of consensus"
repeat_region 56744..56947
/note="L2 repeat: matches 2246..2448 of consensus"
repeat_region 57428..57463
/note="L18 copies 2 mer tt 100% conserved"
misc_feature complement(57559..58119)
/note="match: GSS: Em:A0346372"
repeat_region 58770..59081
/note="Alu repeat: matches 1..309 of consensus"
repeat_region 59703..60000
/note="Alu repeat: matches 1..294 of consensus"
repeat_region 60141..60194
/note="L2 repeat: matches 2652..2706 of consensus"
repeat_region 61142..61188
/note="L2 repeat: matches 2652..2698 of consensus"
repeat_region 61806..61910

```

```

Query Match 2.9%; Score 43; DB 9; Length 105137;
Best Local Similarity 100.0%; Pred. No. 6.8e-13;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1421 CCTTGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 57474 CCTTGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 57432

```

```

RESULT 8
BC022967 BC022967 2871 bp mRNA linear PRI 07-FEB-2002
LOCUS Homo sapiens, clone MGC:30000 IMAGE:5111152, mRNA, complete cds.
DEFINITION BC022967
ACCESSION BC022967.1 GI:18605802
VERSION MGC.
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2871)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs@email.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: amg@bcm.tmc.edu  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 42 Row: m Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction.

```

FEATURES
source
1..2871
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:30000 IMAGE:5111152"
/issue_type="Cervix, Carcinoma"
/clone_lib="NIH-MGC_12"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
7..987
/codon_start=1
/product="Unknown (protein for MGC:30000)"
/protein_id="AAH22867.1"
/db_xref="GI:18603803"
/transcript="MKRAGIGSMRSVVGFLSGRLGDLPLTODFORRLRCRNL
YKDLGHRGCVNAIEFSNMGQVLVSQGDRLVLMHEQALHSRVPDLOKSHS
NIECLAFNSGNTKFEVSGNDKVOYLHDVYLFPHEDAVYGLSVSPNDNFAS
SSDDGVLVMDRESHPGEPCLANVPSAFHVMNPVPEPLRLANSEKVGMDIR
KSSILRYGMLSLQASMSVRFNSNGTOLARLRLPVYDIHSRLPVFPDNGT
FNSCTMKSCFPGRDQHMNMKRRQANSFQFLSLTRNGADVDKLAV"
CDS

```

```

BASE COUNT 802 a 369 c 635 g 865 t
ORIGIN
Query Match 2.7%; Score 40; DB 9; Length 2871;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1424 TTGAGTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 2829 TTGAGTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2868

```

```

RESULT 9
AC084072 AC084072 89181 bp DNA linear HTG 05-FEB-2002
LOCUS Mus musculus clone rp22-51112 strain 12956/SVEYTac, WORKING DRAFT
DEFINITION AC084072
SEQUENCE 17 unordered pieces.
ACCESSION AC084072
VERSION AC084072.6 GI:18497111
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 89181)
AUTHORS Zhu,H., Zullo,S.B. and Roe,B.A.
TITLE Unpublished
JOURNAL Mus musculus BAC Clone rp22-51112
REFERENCE 2 (bases 1 to 89181)
AUTHORS Wu,H., Hu,P., Yang,L., Prescott,A., Zullo,S.J. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (12-OCT-2000) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Feb 5, 2002 this sequence version replaced gi:18390228.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma

```

Center code:UOKNOR

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 17 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 2251: contig of 2251 bp in length  
\* 2252 2351: gap of unknown length  
\* 2352 4397: contig of 2046 bp in length  
\* 4398 4497: gap of unknown length  
\* 4498 6900: contig of 2403 bp in length  
\* 6901 7000: gap of unknown length  
\* 7001 9132: contig of 2132 bp in length  
\* 9133 9232: gap of unknown length  
\* 9233 12520: contig of 3288 bp in length  
\* 12521 12620: gap of unknown length  
\* 12621 14871: contig of 2251 bp in length  
\* 14872 14971: gap of unknown length  
\* 14972 17428: contig of 2457 bp in length  
\* 17429 17528: gap of unknown length  
\* 17529 22646: contig of 5118 bp in length  
\* 22647 22746: gap of unknown length  
\* 22747 26265: contig of 3519 bp in length  
\* 26266 26365: gap of unknown length  
\* 26366 30020: contig of 3655 bp in length  
\* 30021 30120: gap of unknown length  
\* 30121 34472: contig of 4352 bp in length  
\* 34473 34572: gap of unknown length  
\* 34573 40233: contig of 5661 bp in length  
\* 40234 40333: gap of unknown length  
\* 40334 45943: contig of 5610 bp in length  
\* 45944 46043: gap of unknown length  
\* 46044 51644: contig of 5601 bp in length  
\* 51645 51744: gap of unknown length  
\* 51745 60274: contig of 8530 bp in length  
\* 60275 60374: gap of unknown length  
\* 60375 71451: contig of 11077 bp in length  
\* 71452 71551: gap of unknown length  
\* 71552 89181: contig of 17630 bp in length.  
\* Location/Qualifiers  
1. 89181

## FEATURES

source

/organism="Mus musculus"  
/strain="129S6/SvEvTac"  
/db\_xref="taxon:10090"  
/clone="rp22-51112"

BASE COUNT 26250 a 17373 c 16270 g 27668 t 1620 others  
ORIGIN /clone.lib="RC1-22 BAC library"

Query Match 2.7%; Score 40; DB 2; Length 89181;  
Best Local Similarity 100.0%; Pred. No. 3.5e-11;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 TTGAAGTAAAAA 1463

Db 60839 TTGAAGTAAAAA 60878

## RESULT 10

AX185400/c 288 bp DNA linear PAT 06-AUG-2001

LOCUS AX185400 Sequence 1095 from Patent WO0142467.

ACCESSION AX185400

VERSION AX185400.1 GI:15136793

KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE 1 (bases 1 to 288)

AUTHORS Schlegel,R., Deeds,J., Berger,A. and Zhao,X.

TITLE Genes, compositions, kits, and methods for identification,  
assessment, prevention, and therapy of cervical cancer

JOURNAL Patent: WO 0142467-A 1095 14-JUN-2001;  
Millennium Predictive Medicine, Inc. (US)

Location/Qualifiers

## FEATURES

source

1. 288  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 88 a 49 c 42 g 107 t 2 others  
ORIGIN

Query Match 2.7%; Score 39; DB 6; Length 288;  
Best Local Similarity 100.0%; Pred. No. 9.9e-11;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1423 TTGAAGTAAAAA 1461

Db 66 TTGAAGTAAAAA 28

## RESULT 11

AX186826/c 305 bp DNA linear PAT 06-AUG-2001

LOCUS AX186826 Sequence 2521 from Patent WO0142467.

ACCESSION AX186826

VERSION AX186826.1 GI:15138268

KEYWORDS human.  
SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 305)  
AUTHORS Schlegel,R., Deeds,J., Berger,A. and Zhao,X.

TITLE Genes, compositions, kits, and methods for identification,  
assessment, prevention, and therapy of cervical cancer

JOURNAL Patent: WO 0142467-A 2521 14-JUN-2001;  
Millennium Predictive Medicine, Inc. (US)

Location/Qualifiers

## FEATURES

source

1. 305  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 96 a 45 c 51 g 113 t  
ORIGIN

Query Match 2.7%; Score 39; DB 6; Length 305;  
Best Local Similarity 100.0%; Pred. No. 9.9e-11;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1423 TTGAAGTAAAAA 1461

Db 44 TTGAAGTAAAAA 6

## RESULT 12

PFMAL4P4 22448 bp DNA linear HTG 11-AUG-1999

LOCUS PFMAL4P4 Plasmodium falciparum chromosome 4 strain 3D7, \*\*\* SEQUENCING IN  
DEFINITION PROGRAMS \*\*\*; in unordered pieces.

ACCESSION AL035477.5 GI:5731932

VERSION AL035477.5 HTG: HTGS\_PHASE1.

KEYWORDS malaria parasite P. falciparum.  
SOURCE Plasmodium falciparum

ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
REFERENCE 1 (bases 1 to 22448)  
AUTHORS Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quail,M.  
and Barrell,B.

TITLE Direct Submission  
JOURNAL Submitted (24-FEB-1999) P. falciparum Genome Sequencing Consortium,  
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge  
CB10 1SA, UK

## COMMENT

On Aug 12, 1999 this sequence version replaced q1:5531400.  
For more information about this sequence or the Malaria Project,  
see [http://www.sanger.ac.uk/Projects/P\\_falciparum](http://www.sanger.ac.uk/Projects/P_falciparum). IMPORTANT: This  
sequence is unfinished and does not necessarily represent the  
correct sequence. Work on the sequence is in progress and the  
release of this data is based on the understanding that the  
sequence may change as work continues. The sequence may be  
contaminated with foreign sequence from E.coli, yeast, vector,  
phage etc.  
Order of segments is not known: 800 n's separate segments.  
\* NOTE: This is a 'working draft' sequence.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

## FEATURES

Location/Qualifiers

1. 22448  
/organism="Plasmodium falciparum"  
/strain="3D7"  
/db\_xref="taxon:5833"  
/chromosome="4"  
BASE COUNT 80792 a 23258 c 21957 g 75240 t 23201 others  
ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 39; DB 2; Length 22448;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1425 TGAAGTAAAAA  
DB 40702 TGAAGTAAAAA  
1463  
40740

## RESULT 13

BC009510 659 bp mRNA linear PRI 12-JUN-2001  
LOCUS Homo sapiens, Similar to RIKEN cDNA 3110001D03 gene, clone MGC:4730  
IMAGE:353614, mRNA, complete cds.  
ACCESSION BC009510  
VERSION BC009510.1 GI:14550519  
KEYWORDS MGC.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 659)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (25-JUN-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTP

CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia  
Greene, Mark Kettelman and Anuradha Madan  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 11 Row: 1 Column: 6  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis.

FEATURES  
Source Location/Qualifiers  
1. 659  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

## CDS

/clone="MGC:4730 IMAGE:353614"  
/tissue\_type="Lung, small cell carcinoma"  
/clone\_id="NIH\_MGC-7"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
43..393  
/codon\_start=1  
/product="Similar to RIKEN cDNA 3110001D03 gene"  
/protein\_id="AAH09510.1"  
/db\_xref="GI:14550520"  
/translation="MGSRLSQPFESYLTAPGTAAPAKPAPATPGAPTPAEHRL  
KTCMSCPVLGLGIMGAGYVYVWARKPKMGYPSPMTITQVIGLSNGLATWGI  
VWADPKGRKRYRV"  
BASE COUNT 178 a 172 c 174 g 135 t  
ORIGIN

Query Match 2.6%; Score 38; DB 9; Length 659;  
Best Local Similarity 100.0%; Pred. No. 3.8e-10;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAA  
DB 615 GAAGTAAAAA  
1463  
652

RESULT 14  
AX011676 691 bp DNA linear PAT 06-SEP-2000  
LOCUS AX011676  
DEFINITION Sequence 72 from Patent WO955858.  
ACCESSION AX011676  
VERSION AX011676.1 GI:9998200  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 691)  
AUTHORS Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenchal, A. and  
Pillarsky, C.  
TITLE Human nucleic acid sequences obtained from pancreas tumor tissue  
JOURNAL Patent: WO 9955858-A 72 04-NOV-1999;  
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN  
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN  
(DE); PILLARSKY CHRISTIAN (DE)  
FEATURES  
Source Location/Qualifiers  
1. 691  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 192 a 181 c 174 g 144 t  
ORIGIN

Query Match 2.6%; Score 38; DB 6; Length 691;  
Best Local Similarity 100.0%; Pred. No. 3.8e-10;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAA  
DB 631 GAAGTAAAAA  
1463  
668

RESULT 15  
HSM02710 2228 bp mRNA linear PRI 28-JUN-2000  
LOCUS HSM02710  
DEFINITION Homo sapiens mRNA, cDNA DKFP547C146 (from clone DKFP547C146).  
ACCESSION AL359947  
VERSION AL359947.1 GI:8977907  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2228)  
Blum, H., Bauersachs, S., Mewes, H.W., Well, B. and Wilmann, S.

TITLE Direct Submission  
JOURNAL Submitted (15-JUN-2000) MIPS, Am Klopferstritz 18a, D-82152 Martinsried, GERMANY  
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Munich/Germany) within the cDNA sequencing consortium of the German Genome Project.  
This clone (DKFZp547C146) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES  
source Location/Qualifiers  
1..2228  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFZp547C146"  
/clone\_1lb="547 (synonym: hfbrl). Vector psport1; host DH10B; sites NotI + SalI"  
/dev\_stage="fetal"  
/tissue\_type="brain"  
2184

polyA\_site  
BASE COUNT 694 a 403 c 386 g 745 t  
ORIGIN

Query Match 2.6%; Score 38; DB 9; Length 2228;  
Best Local Similarity 100.0%; Pred. No. 4.1e-10;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAA  
Db 2184 GAAGTAAAAA

RESULT 16  
AB042653 2240 bp mRNA linear MAM 23-AUG-2001  
LOCUS Bos taurus mRNA for zona pellucida glycoprotein ZPA, complete cds.  
DEFINITION AB042653  
ACCESSION AB042653.1 GI:12583656  
VERSION  
KEYWORDS  
SOURCE Bos taurus cDNA to mRNA, clone\_1lb:lambda ZAP bovine ovary cDNA (stratagene).  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
1 (bases 1 to 2240)  
Yonezawa, N., Fukui, N., Kuno, M., Shinoda, M., Goko, S., Mitsui, S. and Nakano, M.  
Molecular cloning of bovine zona pellucida glycoproteins ZPA and ZPB and analysis for sperm-binding component of the zona  
European Journal of Biochemistry / FEBS. 268 (12), 3587-3594 (2001)  
21315847  
2 (bases 1 to 2240)  
Nakano, M. and Yonezawa, N.  
Direct Submission  
Submitted (15-MAY-2000) Minoru Nakano, Chiba University, Department of Chemistry, Inage-ku, 1-33 Yayoi-cho, Chiba 263-8522, Japan (E-mail: minakano@sci.chem.s.chiba-u.ac.jp, Tel:81-43-290-2794, Fax:81-43-290-2874)

FEATURES  
source Location/Qualifiers  
1..2240  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_1lb="lambda ZAP bovine ovary cDNA (stratagene)"  
30..2171  
/gene="ZPA"  
30..2171  
/gene="ZPA"  
/codon\_start=1  
/product="zona pellucida glycoprotein ZPA"

/protein\_id="BAB21482.1"  
/db\_xref="GI:12583657"  
/translation="MACRORGDGSRPSSWEPADRSFPLSTLTYSNIDVNOIDPA  
FEGLVATCTENRMVVEFKRTLNKQIOMHSYNDLCLMKNCYVYIDPRLTLKAYESC  
TKRVLGQHMRTTENDNTAHROKTVLTHVSCPWQAGRHDSGTICKDPSFPE  
HFFPLADDTGPRKPMGMVTVVDGEAQLTVQELAQDYLINLLNOKMSIOVLFH  
ATGVTHSQGSHIMVPLKTHVSPGQTILLSSRLICADPPVCNATHMTLTPERP  
GRKSVSENNKINAVNOLNMSGIYVIELGLRHLHSTLTKFSEKCLPYOFLSL  
KLTFYOLETYSMTVYIPECVESTYSIYSGELCTODGFMDEVYRHOTKPALNDLTLR  
VDSGSCQPTKAPROGLYKPHIPLNGCCTRRKFPNGKVIYENEIHALMADLPSTISR  
DSERHRTVRCITYSSNNLTNNVESLPPPAVSYPGPAITLQYTPDPSYLOPGDND  
YPVVRYLQPTILEVRLNRPDKNTEFVSDLPKMATSTNDPASTLNNNTIYDGCYVND  
NHRTHFHVGSVAVPNHYORFAKTEFVSDLPKMATSTNDPASTLNNNTIYDGCYVND  
CSASCLVSSRSRRATGATEEERKMIVSLPGLILLSDGSSFRDAVDSKGHTSGYAAK  
TWAVVALAGVAATLSTLSYLRKKRIYVNLN"

BASE COUNT 635 a 536 c 485 g 584 t  
ORIGIN

Query Match 2.6%; Score 38; DB 4; Length 2240;  
Best Local Similarity 100.0%; Pred. No. 4.1e-10;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAA  
Db 2193 GAAGTAAAAA

RESULT 17  
BC001537 2291 bp mRNA linear PRI 22-OCT-2001  
LOCUS Homo sapiens, clone MGC:786 IMAGE:2987938, mRNA, complete cds.  
DEFINITION BC001537  
ACCESSION BC001537.1 GI:16306716  
VERSION  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 2291)  
Strausberg, R.  
Direct Submission  
Submitted (21-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
DNA Library Preparation: Rubin laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan

REMARK  
COMMENT  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 3 Row: 3 Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7657474.

FEATURES  
source Location/Qualifiers  
1..2291  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:786 IMAGE:2987938"  
/tissue\_type="Colon, adenocarcinoma"  
/clone\_1lb="NIH\_MGC\_15"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
51..1820  
/codon\_start=1

/product="Unknown (protein for MG:786)"  
/protein\_id="AAH01537.1"  
/db\_xref="GI:16306717"  
/translation="MAADGDSLPYPIAVLIDELRNEVQRLNSIKLSTYALGV  
ETRELLPFLITDTITDEDEVLALAEBOJGFTTLVGSPVHCLPLESLAVEET  
VVRDAVSLRLAISHSHPDLEAFVPLVRKLAGDWFSTRSACGLFVYCYRVS  
AVKALRYFRNLCSDDTPMVRRAASKLGFAKLELDNKSEIIPMSNLASDEOD  
SVRLAVACVNIQOLLPOEDLEALVMTLRQAEKSWRVYVADKETEOKAVGP  
ETTKDVLPAFONLMDCEAEVRAASHKVEPCNLSADCREVYMQIIPCKEY  
SPANHVSALASVIMGSLPITGKNTLEHLLPLFLADKCECEVRNLITSLDCVN  
EYICGRQSLQSLPAIVEAEADAKRVRALITETVPLLAGDQVPEDEKINSLCMV  
LVDHYAIRKATSNLKKLVERFKEMAHATIPKIVLMSGDPMPLHMTLFCINVL  
SEVCGDITTKHMLPFLVLRMAGDPVAVNRFVNAKSLOKIGPILDNSTLOSEVRLKLEK  
LTQDDVDVKFAQALTVLSIA"

BASE COUNT 490 a 682 c 642 g 477 t

Query Match 2.6%; Score 38; DB 9; Length 2291;  
Best Local Similarity 100.0%; Pred. No. 4.1e-10;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1426 GAAGTAAAAA... 1463  
Db 2252 GAAGTAAAAA... 2289

RESULT 18  
AY069109 3315 bp mRNA linear INV 17-DEC-2001  
LOCUS Drosophila melanogaster GH12580 full length cDNA.  
DEFINITION AY069109  
ACCESSION AY069109.1 GI:17861553  
KEYWORDS FLI\_CDNA.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 3315)  
Stapleton, M., Brockstein, P., Hong, L., Abghayani, A., Carlson, J.,  
Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,  
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,  
Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,  
Yu, C., Lewis, S. E., Rubin, G. M. and Celinker, S.  
Direct Submission  
Submitted (10-DEC-2001) Berkeley Drosophila Genome Project,  
Lawrence Berkeley National Laboratory, One Cyclotron Road,  
Berkeley, CA 94720, USA  
COMMENT  
Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720  
This clone was sequenced as part of a high-throughput process to  
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
Science 2000). The sequence has been subjected to integrity checks  
for sequence accuracy, presence of a polyA tail and contiguity  
within 100 kb in the genome. Thus we believe the sequence to  
reflect accurately this particular cDNA clone. However, there are  
artifacts associated with the generation of cDNA clones that may  
have not been detected in our initial analyses such as internal  
priming, priming from contaminating genomic DNA, retained introns  
due to reverse transcription of unspliced precursor RNAs, and  
reverse transcriptase errors that result in single base changes.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our web site  
(http://fruitfly.berkeley.edu) or send email to  
cna@fruitfly.berkeley.edu

FEATURES  
SOURCE  
1..3315  
/organism="Drosophila melanogaster"  
/strain="y; cn bw sp"  
/db\_xref="taxon:7227"  
/map="86C2-86C2"

gene  
1..3315  
/gene="Skeletor"  
/note="alignment with genomic scaffold AE003688"  
/db\_xref="FLYBASE:FBgn0020517"  
1933..3054  
CDS  
/gene="Skeletor"  
/note="Longest ORF"  
/codon\_start=1  
/db\_xref="FLYBASE:FBgn0020517"  
/product="GH12580p"  
/protein\_id="AAL39254.1"  
/db\_xref="GI:17861554"  
/translation="MRPVRVNGGRKPDVTYVESGFRIVTIDGTGVLDLPREITIDYVHR  
RBPGRFTIDVMTDTLFLAQQGSGSTOSFEFMTPTSPDSTNATKVLRLVNVKESV  
TASALRTPSALFHALPSASELTKPTLDELFAEDLNEBELMEKPFVADVESLEET  
KDAVYTTINIPRNTTKKPDPLLEDLFGDEBELYADELEMDRVAAMARIIDTY  
YLPDNRKIPDTRVPSGALYTFDGKSVDSSTLYLPKLDAPDANVHQRHAQYGLPL  
EOLVRTTPORGVYRGELPOEFGTEPOPVSEVSHAPPSRTTPVVFSSSGSTIYPRSS  
STGASVTVSSSSPLSSSLRPSTKIQLLKPEGRRA"

BASE COUNT 880 a 992 c 812 g 631 t

Query Match 2.6%; Score 38; DB 3; Length 3315;  
Best Local Similarity 100.0%; Pred. No. 4.1e-10;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1426 GAAGTAAAAA... 1463  
Db 3274 GAAGTAAAAA... 3311

RESULT 19  
AC105102/c 67632 bp DNA linear HTG 23-DEC-2001  
LOCUS Homo sapiens chromosome 17 clone RP11-316K20 map 17, LOW-PASS  
DEFINITION AC105102  
ACCESSION AC105102  
VERSION AC105102.1 GI:17977598  
KEYWORDS HTG; HTGS\_PHASE0.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 67632)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 17, clone RP11-316K20  
Unpublished  
2 (bases 1 to 67632)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,  
Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,  
Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,  
Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N.,  
Hagob, B., Headford, A., Horton, L., Hulme, W., Illiev, I., Johnson, R.,  
Jones, C., Kamt, A., Karatas, A., Kells, C., Larocque, K.,  
Lamazares, R., Landers, T., Jehoczky, J., Levine, R., Liu, G.,  
Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,  
McCarthy, M., McGowan, P., McKernan, K., McPheters, R., Meldrim, J.,  
Meneus, L., Mihova, T., Mianga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Struss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Triggilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (23-DEC-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIRB

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L22879

Center clone name: 316\_K\_20

-----  
\* NOTE: This record contains 80 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 748: contig of 748 bp in length  
\* 749 848: gap of 100 bp  
\* 849 1565: contig of 717 bp in length  
\* 1566 1665: gap of 100 bp  
\* 1666 2391: contig of 726 bp in length  
\* 2392 2491: gap of 100 bp  
\* 2492 3242: contig of 751 bp in length  
\* 3243 3342: gap of 100 bp  
\* 3343 4107: contig of 765 bp in length  
\* 4108 4207: gap of 100 bp  
\* 4208 4975: contig of 768 bp in length  
\* 4976 5075: gap of 100 bp  
\* 5076 5833: contig of 758 bp in length  
\* 5834 5933: gap of 100 bp  
\* 5934 6686: contig of 753 bp in length  
\* 6687 6786: gap of 100 bp  
\* 6787 7545: contig of 759 bp in length  
\* 7546 7645: gap of 100 bp  
\* 7646 8374: contig of 729 bp in length  
\* 8375 8474: gap of 100 bp  
\* 8475 9218: contig of 744 bp in length  
\* 9219 9318: gap of 100 bp  
\* 9319 10059: contig of 741 bp in length  
\* 10060 10159: gap of 100 bp  
\* 10160 10901: contig of 742 bp in length  
\* 10902 11001: gap of 100 bp  
\* 11002 11748: contig of 747 bp in length  
\* 11749 11848: gap of 100 bp  
\* 11849 12609: contig of 761 bp in length  
\* 12610 12709: gap of 100 bp  
\* 12710 13472: contig of 763 bp in length  
\* 13473 13572: gap of 100 bp  
\* 13573 14310: contig of 738 bp in length  
\* 14311 14410: gap of 100 bp  
\* 14411 15163: contig of 753 bp in length  
\* 15164 15263: gap of 100 bp  
\* 15264 16023: contig of 760 bp in length  
\* 16024 16123: gap of 100 bp  
\* 16124 16872: contig of 749 bp in length  
\* 16873 16972: gap of 100 bp  
\* 16973 17717: contig of 745 bp in length  
\* 17718 17817: gap of 100 bp  
\* 17818 18561: contig of 744 bp in length  
\* 18562 18661: gap of 100 bp  
\* 18662 19402: contig of 741 bp in length  
\* 19403 19502: gap of 100 bp  
\* 19503 20217: contig of 715 bp in length  
\* 20218 20317: gap of 100 bp  
\* 20318 21054: contig of 737 bp in length  
\* 21055 21154: gap of 100 bp

\* 21155 21911: contig of 757 bp in length  
\* 21912 22011: gap of 100 bp  
\* 22012 22770: contig of 759 bp in length  
\* 22771 22870: gap of 100 bp  
\* 22871 23627: contig of 757 bp in length  
\* 23628 23727: gap of 100 bp  
\* 23728 24480: contig of 753 bp in length  
\* 24481 24580: gap of 100 bp  
\* 24581 25327: contig of 747 bp in length  
\* 25328 25427: gap of 100 bp  
\* 25428 26179: contig of 752 bp in length  
\* 26180 26279: gap of 100 bp  
\* 26280 27038: contig of 759 bp in length  
\* 27039 27138: gap of 100 bp  
\* 27139 27886: contig of 748 bp in length  
\* 27887 27986: gap of 100 bp  
\* 27987 28738: contig of 752 bp in length  
\* 28739 28838: gap of 100 bp  
\* 28839 29582: contig of 744 bp in length  
\* 29583 29682: gap of 100 bp  
\* 29683 30427: contig of 745 bp in length  
\* 30428 30527: gap of 100 bp  
\* 30528 31245: contig of 718 bp in length  
\* 31246 31345: gap of 100 bp  
\* 31346 32107: contig of 762 bp in length  
\* 32108 32207: gap of 100 bp  
\* 32208 32972: contig of 765 bp in length  
\* 32973 33072: gap of 100 bp  
\* 33073 33828: contig of 756 bp in length  
\* 33829 33928: gap of 100 bp  
\* 33929 34693: contig of 765 bp in length  
\* 34694 34793: gap of 100 bp  
\* 34794 35535: contig of 742 bp in length  
\* 35536 35635: gap of 100 bp  
\* 35636 36381: contig of 746 bp in length  
\* 36382 36481: gap of 100 bp  
\* 36482 37177: contig of 696 bp in length  
\* 37178 37277: gap of 100 bp  
\* 37278 38004: contig of 727 bp in length  
\* 38005 38104: gap of 100 bp  
\* 38105 38821: contig of 717 bp in length  
\* 38822 38921: gap of 100 bp  
\* 38922 39675: contig of 754 bp in length  
\* 39676 39775: gap of 100 bp  
\* 39776 40525: contig of 750 bp in length  
\* 40526 40625: gap of 100 bp  
\* 40626 41380: contig of 755 bp in length  
\* 41381 41480: gap of 100 bp  
\* 41481 42244: contig of 764 bp in length  
\* 42245 42344: gap of 100 bp  
\* 42345 43098: contig of 754 bp in length  
\* 43099 43198: gap of 100 bp  
\* 43199 43944: contig of 746 bp in length  
\* 43945 44044: gap of 100 bp  
\* 44045 44800: contig of 756 bp in length  
\* 44801 44900: gap of 100 bp  
\* 44901 45660: contig of 760 bp in length  
\* 45661 45760: gap of 100 bp  
\* 45761 46514: contig of 754 bp in length  
\* 46515 46614: gap of 100 bp  
\* 46615 47355: contig of 741 bp in length  
\* 47356 47455: gap of 100 bp  
\* 47456 48214: contig of 759 bp in length  
\* 48215 48314: gap of 100 bp  
\* 48315 49059: contig of 745 bp in length  
\* 49060 49159: gap of 100 bp  
\* 49160 49919: contig of 760 bp in length  
\* 49920 50019: gap of 100 bp  
\* 50020 50725: contig of 706 bp in length  
\* 50726 50825: gap of 100 bp  
\* 50826 51581: contig of 756 bp in length  
\* 51582 51681: gap of 100 bp  
\* 51682 52420: contig of 739 bp in length



```

* 52421 52520: gap of 100 bp
* 52521 53253: contig of 733 bp in length
* 53254 53353: gap of 100 bp
* 53354 54092: contig of 739 bp in length
* 54093 54192: gap of 100 bp
* 54193 54927: contig of 735 bp in length
* 54928 55027: gap of 100 bp
* 55028 55724: contig of 697 bp in length
* 55725 55824: gap of 100 bp
* 55825 56595: contig of 771 bp in length
* 56596 56695: gap of 100 bp
* 56696 57447: contig of 752 bp in length
* 57448 57547: gap of 100 bp
* 57548 58290: contig of 743 bp in length

```

Query Match 2.6%: Score 38; DB 2; Length 67632;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-10;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 DB 35063 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 35026

RESULT 20  
 PPMALIP2/c 98734 bp DNA linear HTG 11-AUG-1999  
 LOCUS Plasmodium falciparum chromosome 1 strain 3D7, \*\*\* SEQUENCING IN  
 DEFINITION PROGRESS \*\*\*, in unordered pieces.  
 ACCESSION AL031745  
 VERSION AL031745.7 GI:5731893  
 KEYWORDS HTG; HTGS\_PHASE1.  
 SOURCE malaria parasite P. falciparum.  
 ORGANISM Plasmodium falciparum  
 Eukaryotes; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 1 (bases 1 to 98734)  
 AUTHORS Bowman, S., Churcher, C., Harris, B., Harris, D., Lawson, D., Quail, M.  
 and Barrell, B.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-SEP-1998) P. falciparum Genome Sequencing Consortium,  
 The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge  
 CB10 1SA, UK  
 On Aug 12, 1999 this sequence version replaced gi:5706497.

For more information about this sequence or the Malaria Project,  
 see <http://www.sanger.ac.uk/projects/P-falciparum>. IMPORTANT: This  
 sequence is unfinished and does not necessarily represent the  
 correct sequence. Work on the sequence is in progress and the  
 release of this data is based on the understanding that the  
 sequence may change as work continues. The sequence may be  
 contaminated with foreign sequence from E.coli, yeast, vector,  
 phage etc.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

FEATURES  
 source 1..98734  
 /organism="Plasmodium falciparum"  
 /strain="3D7"  
 /db\_xref="taxon:5833"  
 /chromosome="1"

BASE COUNT 37227 a 8619 c 9384 g 35504 t 8000 others  
 ORIGIN

Query Match 2.6%: Score 38; DB 2; Length 98734;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-10;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 DB 16502 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 16465

RESULT 21  
 AL354982/c 119077 bp DNA linear PRI 16-OCT-2001  
 LOCUS Human DNA sequence from clone RP11-473H13 on chromosome 9, complete  
 DEFINITION sequence.  
 ACCESSION AL354982  
 VERSION AL354982.12 GI:16214577  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 119077)

REFERENCE  
 AUTHORS Direct Submission  
 TITLE Submitted (16-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 JOURNAL CB10 1SA, UK. E-mail enquiries: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk) clone  
 requests: [clonerequests@sanger.ac.uk](mailto:clonerequests@sanger.ac.uk)  
 On Oct 17, 2001 this sequence version replaced gi:15131224.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em., EMBL; Sw.,  
 SWISSPROT; Tr., TrEMBL; Wp., WormPEP; information on the WormPEP  
 database can be found at  
<http://www.sanger.ac.uk/projects/C-elegans/wormpep> This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping  
 Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr9>  
 RP11-473H13 is from the library RCT-11.2 constructed by the group  
 of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6  
 IMPORTANT: This sequence is not the entire insert of clone  
 RP11-473H13. It may be shorter because we sequence overlapping  
 sections only once, except for a short overlap.  
 The true left end of clone RP11-410X21 is at 117078 in this  
 sequence. The true right end of clone RP11-427J11 is at 2000 in  
 this sequence.

FEATURES  
 source 1..119077  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="9"  
 /clone="RP11-473H13"  
 /clone\_1b="RPCT-11.2"  
 /clone\_1lb="RPCT-11.2"

BASE COUNT 36688 a 22359 c 21866 g 37984 t  
 ORIGIN

Query Match 2.6%: Score 38; DB 9; Length 119077;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-10;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 DB 83006 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 82969

RESULT 22  
 HS199L16/c

LOCUS H5199L16 134963 bp DNA linear PRI 23-NOV-1999  
 DEFINITION Human DNA sequence from clone 199L16 on chromosome Xq22.1-22.3  
 Contains a 40S Ribosomal Protein S29 Like (pseudor) gene, an EST,  
 STSs, GSSs and tcta and tatg repeat polymorphisms, complete  
 sequence.  
 ACCESSION AL022151 GI:3790137  
 VERSION AL022151.1  
 KEYWORDS RTG: 40S Ribosomal Protein S29; tatg repeat polymorphisms; tcta  
 repeat polymorphisms.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 134963)  
 AUTHORS Errington,H.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-OCT-1998) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 COMMENT requests: clonerequest@sanger.ac.uk  
 On Oct 26, 1998 this sequence version replaced gi:3646052.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence is the entire insert of clone 199L16. The true right  
 end of clone 468H17 (AL031114) is at 26354 in this sequence.  
 This sequence has been finished according to sequence map criteria  
 as follows. An attempt is made to resolve all sequencing problems,  
 such as compressions and repeats, but not necessarily within known  
 annotated human repeat sequence elements (e.g. Alu). Where the  
 sequence is ambiguous, there is an annotation using the 'unsure'  
 feature key.  
 This sequence was generated from part of bacterial clone contigs of  
 human chromosome X, constructed by the Sanger Centre Chromosome X  
 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chrx  
 199L16 is from the library RPCI constructed at the Roswell Park  
 Cancer Institute by the group of Pieter de Jong. For further  
 details see http://hacpac.med.buffalo.edu/VECTOR: pcrpnc2.  
 FEATURES  
 source  
 1..134963  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 /map="q22.1-22.3"  
 /clone="RPI-199L16"  
 /clone\_lib="RPCI-1"  
 1..237  
 /note="AluX repeat: matches 1..237 of consensus"  
 239..5892  
 /note="L1M1 repeat: matches -1390..4370 of consensus"  
 6073..6262  
 /note="L1MC/D repeat: matches 5493..5700 of consensus"  
 6264..6351  
 /note="AluSg/X repeat: matches 210..297 of consensus"  
 6355..8296  
 /note="L1MB1 repeat: matches 3986..5896 of consensus"  
 8291..8555  
 /note="L1PB repeat: matches 257..523 of consensus"  
 8559..8584  
 /note="L1PB repeat: matches 1990..2015 of consensus"  
 8585..8876  
 /note="L1M4C repeat: matches 1990..2015 of consensus"  
 8877..9023  
 /note="L1PA8 repeat: matches 5874..6162 of consensus"  
 9024..9128  
 /note="L1M4C repeat: matches 1869..1991 of consensus"  
 9122..9374  
 /note="R1GGER1 repeat: matches 1..99 of consensus"  
 9375..9637  
 /note="R1GGER1 repeat: matches 1823..2069 of consensus"  
 9375..9637  
 /note="AluSp repeat: matches 50..313 of consensus"

repeat\_region 9638..9659  
 /note="R1GGER1 repeat: matches 2069..2090 of consensus"  
 repeat\_region 9696..11174  
 /note="L1PA12 repeat: matches -1414..1133 of consensus"  
 mRNA complement(11400..11692)  
 /gene="dJ199L16.1"  
 /note="match: cDNAs X59051 U66372 U14973 L31609 L31610;  
 could be a pseudogene"  
 /product="dJ199L16.1 (40S Ribosomal Protein S29 LIKE)"  
 gene evidence-not-experimental  
 complement(11400..11692)  
 /gene="dJ199L16.1"  
 /complement(11400)  
 /gene="dJ199L16.1"  
 /complement(11406..11411)  
 /gene="dJ199L16.1"  
 complement(11482..>11634)  
 /gene="dJ199L16.1"  
 /note="match: proteins P14041 P54110 P30054 O15619 P26816  
 O28368 O05635 P41057 P90983 O26125 P41058; could be a  
 pseudogene"  
 /codon\_start=1  
 /evidence-not-experimental  
 /product="dJ199L16.1 (40S Ribosomal Protein S29 LIKE)"  
 /protein\_id="CAI8145.1"  
 /db\_xref="GI:3790138"  
 /translation="FSQGSNSCQVYSNMHGLIKYGLNVCQCFHOYLDIGFIKLDY  
 SKKILIK"  
 repeat\_region 11801..11900  
 /note="L1MC/D repeat: matches 5509..5619 of consensus"  
 repeat\_region 12006..12191  
 /note="L1M4 repeat: matches 2657..3049 of consensus"  
 repeat\_region 12200..12599  
 /note="MLT2B repeat: matches 1..400 of consensus"  
 12566..12883  
 /note="match: STS G08105"  
 12595..12630  
 /note="9 copies of tatg 100% conserved; differs from  
 G08105"  
 repeat\_region 12606..12657  
 /note="26 copies 2 mer at 75% conserved"  
 12633..12676  
 /note="11 copies of tcta 100% conserved; differs from  
 G08105"  
 repeat\_region 12660..12726  
 /note="MLT2B repeat: matches 384..448 of consensus"  
 13405..13810  
 /note="L2 repeat: matches 1259..1672 of consensus"  
 repeat\_region 14054..14262  
 /note="WIR repeat: matches 20..216 of consensus"  
 complement(14545..14834)  
 /note="match: EST AA328142"  
 14837..15005  
 /note="MER20 repeat: matches 4..174 of consensus"  
 15164..15275  
 /note="L2 repeat: matches 2626..2750 of consensus"  
 15342..15466  
 /note="WIR repeat: matches 13..139 of consensus"  
 15952..16120  
 /note="MER5A repeat: matches 2..154 of consensus"  
 16121..16317  
 /note="MER45 repeat: matches 2..174 of consensus"  
 16318..16351  
 /note="MER5A repeat: matches 154..189 of consensus"  
 16469..20086  
 /note="L1PA2 repeat: matches 2519..6144 of consensus"  
 20102..20214  
 /note="L2 repeat: matches 1991..2102 of consensus"  
 complement(22045..22272)  
 /note="match: STS L42655"  
 24543..24722  
 /note="MLT1A1 repeat: matches 172..365 of consensus"  
 24726..24874

```

repeat_region      /note="L1R16 repeat: matches 314. .464 of consensus"
25012. .25360
/note="L1R16 repeat: matches 19. .365 of consensus"
25361. .25643
/note="MER61C repeat: matches 134. .431 of consensus"
25691. .25878
/note="MER61B repeat: matches 1. .178 of consensus"
26111. .26322
/note="L1R13 repeat: matches 14. .260 of consensus"
27283. .27498
/note="L1R13 repeat: matches 5831. .6046 of consensus"
27499. .27555
/note="MER4D repeat: matches 410. .466 of consensus"
27659. .27810
/note="MER4D repeat: matches 522. .687 of consensus"
27847. .28049
/note="L1R29 repeat: matches 1. .202 of consensus"
28071. .28091
/note="L1M4A repeat: matches 6038. .6057 of consensus"
28351. .28644
/note="L1M4 repeat: matches 1. .293 of consensus"
28708. .28939
/note="L1P16 repeat: matches 5919. .6156 of consensus"
28963. .29076
/note="L1M4 repeat: matches 7857. .7975 of consensus"
29097. .29644
/note="L1M4 repeat: matches 7079. .7693 of consensus"
30797. .30977
/note="L1R8 repeat: matches 1. .182 of consensus"
30876. .31000
/note="MER8 repeat: matches 66. .190 of consensus"
31031. .31323
/note="L1R8 repeat: matches 397. .691 of consensus"
31418. .31451
/note="L1R17 repeat: matches 2 mer tt 82% conserved"
31454. .32540
/note="L1P2 repeat: matches 5058. .6144 of consensus"
32543. .32665
/note="L1M repeat: matches 5260. .5389 of consensus"
32879. .33080
/note="L1 repeat: matches 4085. .4292 of consensus"
33648. .33808
/note="match: STS AF020146"
34410. .34512
/note="L1P10 repeat: matches 6063. .6165 of consensus"
34639. .34702
/note="L1R32 repeat: matches 2 mer ga 70% conserved"
34839. .35073
/note="L1R repeat: matches 45. .256 of consensus"
34961. .35083
/note="L1R2 repeat: matches 2626. .2748 of consensus"
35350. .35423
/note="L1R repeat: matches 72. .142 of consensus"
37390. .37568
repeat_region

```

Query Match 2.6%; Score 38; DB 9; Length 134963;  
 Best Local Similarity 100.0%; Pred.No.: 4.9e-10;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yr 1426 GAGGTAAAAA  
 DB 132649 GAGGTAAAAA  
 AC012032  
 LOCUS AC012032 145069 bp DNA linear HTG 16-JAN-2002  
 DEFINITION Homo sapiens chromosome 3 clone RP11-630D12, WORKING DRAFT  
 SEQUENCE, 3 unordered pieces.  
 AC012032  
 AC012032  
 AC012032  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 KEYWORDS human.

## ORGANISM

Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

## AUTHORS

Mazny,D.M., Adams,C., Adio-Odola,B., Al-Isman,F.R., Allen,C.,  
 Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbata,J.,  
 Benton,D., Bimaga,K., Blankenburg,K., Bonin,D., Bouck,J.,  
 Bowls,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Bulay,C.,  
 Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
 Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
 Davy,Carroll,L., Dederich,D.A., Delaney,K.R., Deigard,O.,  
 Demn,A.L., Ding,Y., Dinn,H.H., Douthwaite,K.J., Draper,H.,  
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
 Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A.,  
 Hernandez,J., Hernandez,O., Hodgson,A., Hughes,M., Holloway,C.,  
 Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,  
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Joliver,S.,  
 Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
 Kovar,C., Kralovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
 Louisege,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
 Martinez,E., Massey,E., Mathiney,E., McLeod,M.P., Meador,M.,  
 Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,  
 Oguh,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojucokan,I., Rolfe,M.,  
 Riz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N.,  
 Sisson,I., Sodergren,E., Soneike,T., Sparks,A., Stanley,H.,  
 Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,  
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
 Wellington,S., Williams,G., Williamson,A., Wlezyk,R., Wooden,S.,  
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstock,G. and Gibbs,R.

Direct Submission  
 Unpublished  
 2 (bases 1 to 145069)  
 Worley,K.C.

Direct Submission  
 Submitted (19-OCT-1999) Human genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jan 16, 2002 this sequence version replaced gi:1723148.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: HMO  
 Center clone name: RP11-630D12  
 ----- Summary Statistics  
 Sequencing vector: M13; I08821  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329first call to  
 findPhrapList  
 Consensus quality: 148653 bases at least Q40  
 Consensus quality: 154105 bases at least Q30  
 Consensus quality: 158675 bases at least Q20  
 Estimated insert size: 153129; sum-of-contigs estimation  
 Estimated insert size: 208334; agarose-fp estimation  
 Quality coverage: 6.5x in Q20 bases; agarose-fp estimation  
 Quality coverage: 8.9x in Q20 bases; sum-of-contigs estimation



619	<800	3279	3420	6680	6566
2743	2750	899	926	2785	2791
2449	2497	656	<800	2400	2445
4327	4297	3357	3502	1764	1791
166	<800	507	<800	3064	3285
3226	3295	1994	2075	3900	3845
1497	1475	527	<800	4935	4848
3208	3295	4216	4360	5344	5271
2188	2164	1648	1642	999	1018
239	<800	4284	4360	2222	2237
586	<800	289	<800	1819	1791
1874	1823	8248	8328	2741	2791
652	<800	1406	1368	5859	5782
274	<800	1861	1943	3444	3440
772	861	11358	11428	1019	1018
154	<800	448	<800	2749	2791
1494	1475	1538	1502	2790	2791
311	<800	297	<800	1211	1166
1395	1386	1085	1077	3377	3440
1156	1189	2526	2681	2493	2445
170	<800	6144	6200	1159	1166
1244	1189	3987	3730	373	<800
239	<800	6763	6900	1490	1466
4367	4297	5304	5584	2457	2445
1208	1189	3197	3332	1154	1166
2660	2750	2739	2864	1823	1791
2142	2164	931	926	16	<800
835	861	662	<800	6749	6758
79	<800	4170	4360	443	<800
5059	4966	4542	4642	387	<800
10502	10502	1013	1019	4638	4604
4292	4297	5739	5806	11487	11490
1179	1189	169	<800	6981	6948
1344	1386	401	<800	5635	5567
1649	1619	43	<800	436	<800
2141	2164	1482	1502	3315	3440

Query Match	2.68;	Score 38;	DB 9;	Length 165894;
Best Local Similarity	100.0%;	Pred. No. 5e-10;		
Matches 38;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Q7 1420 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
| | | | | | | | | | | | | | | | | | | | | |  
Db 143412 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 14375

## RESULT 25

AC027646

LOCUS

DEPT 11110

ACCESSION

VERSION

## KEYWORDS

**SOURCE**

## ORGANISMS

## REFERENCE

## AUTHORS

ATTN: E

FILE  
JOURNAL

## REFERENCE

## AUTHORS

TITLE \_\_\_\_\_  
 TOPIC \_\_\_\_\_

JOURNAL

COMMENT

1424	1386	5453	5390	3767	3687
653	<800	10759	10721	4845	4848
631	<800	1633	1642	2909	3054
2752	2750	2970	3097	359	<800
2681	2750	6554	6426	1186	1166
19	<800	3438	3579	853	872
975	975			1871	1791
4157	4086				

ch 2.6% Score 38: DB 9: Length 165894;  
1 Similarity 100.0% Pred. No. 5e-10;  
38: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

AAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 143375

AC027646 196560 bp DNA linear HTG 25-JAN-2002  
Mus musculus clone RP23-18J14 strain C57BL6/J, WORKING DRAFT  
SEQUENCE, 21 unordered pieces.  
AC027646  
AC027646.9 GI:14149542  
HTG: HTGS\_PHASE1, HTGS\_DRAFT.  
house mouse.  
Mus musculus  
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;  
Mammalia: Eutheria: Rodentia: Sciuromorphi: Muridae: Murinae: Mus.  
1 (bases 1 to 196560)  
Han,J., Montgomery,K.T., Grills,G., Lee,E., Long,J., Pomerantz,R.,  
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,  
Gordon,M., Goltz,J.S. and Kucherlapati,R.  
Submitted (31-MAR-2000) Department of Molecular Genetics, Albert  
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,  
Bronx, NY 10461, USA  
On May 19, 2001 this sequence version replaced gi:14091719.  
Genome Center  
Center: Harvard Partners Genome Center  
Center Code: HPGC  
Web site:  
http://wchanning.bwh.harvard.edu:9088/hpcg/jsp/hpcg/sequence/mouse.html  
Contact: gnikm@capred.bwh.harvard.edu

Summary Statistics  
Center project name: ABE  
Sequencing vector: pUC18; L08752  
Chemistry: Dye-terminator Big Dye; 100%  
\*Consensus quality: 190345 at least Q20  
\*Consensus quality: 189126 at least Q30  
\*Consensus quality: 187254 at least Q40  
Estimated insert size: agarose-ftp - N/A  
\*Estimated insert size: 196160 - sum-of-contigs  
Quality coverage: agarose-ftp - N/A  
Quality coverage: 8.6 x in Q20 bases; sum-of-contigs estimation

\* NOTE: This is a 'working draft' sequence. It currently

*	1	51804:	contig of 51804 bp in length
*	51805	518204:	gap of unknown length
*	51825	90706:	contig of 38882 bp in length
*	90707	90726:	gap of unknown length
*	90727	123879:	contig of 33153 bp in length
*	123880	123899:	gap of unknown length
*	123900	143393:	contig of 13394 bp in length
*	143294	143313:	gap of unknown length
*	143314	163801:	contig of 20488 bp in length
*	163802	163821:	gap of unknown length
*	163822	173127:	contig of 9306 bp in length
*	173128	173147:	gap of unknown length
*	173148	182893:	contig of 9546 bp in length
*	182694	182713:	gap of unknown length
*	182714	183958:	contig of 1245 bp in length
*	183959	183978:	gap of unknown length
*	183979	185019:	contig of 1041 bp in length
*	185020	185039:	gap of unknown length
*	185040	186528:	contig of 1489 bp in length
*	186529	186548:	gap of unknown length
*	186549	187616:	contig of 1068 bp in length
*	187617	187636:	gap of unknown length
*	187637	187921:	contig of 285 bp in length
*	187922	187941:	gap of unknown length
*	188538	188558:	contig of 597 bp in length
*	188539	188558:	gap of unknown length
*	188559	190007:	contig of 1449 bp in length
*	190008	190027:	gap of unknown length
*	190028	190585:	contig of 558 bp in length
*	190586	190605:	gap of unknown length
*	190606	192178:	contig of 1573 bp in length
*	192179	192198:	gap of unknown length
*	192199	193302:	contig of 1104 bp in length
*	193303	193322:	gap of unknown length
*	193323	194502:	contig of 1180 bp in length
*	194503	194522:	gap of unknown length
*	194523	195116:	contig of 594 bp in length
*	195117	195136:	gap of unknown length
*	195137	195631:	contig of 545 bp in length
*	195632	195701:	gap of unknown length
*	195702	195650:	contig of 859 bp in length

```
misc_feature
1. "assembly
/ote="assembly_name:ContI221"
151825. 90706
/ote="assembly_name:ContI220
clone_end:r7
vector_side:left"
90727.123879
/ote="assembly_name:ContI219
123900. 143393
/ote="assembly_name:ContI218
143314. 163801
/ote="assembly_name:ContI217
163822. 1713127
/ote="assembly_name:ContI216
clone_end:Sp6
vector_side:right"
173148. 182693
/ote="assembly_name:ContI215
```

Query Match	2.6%;	Score 38;	DB 2;	Length 196560;
Best Local Similarity	100.0%;	Pred. No. 5e-10;		
Matches	38;	Conservative	0;	Mismatches 0;
				Indels 0;
				Caps 0;
QY	1426	GAAGTAAAAAAAAAAAAAAAAAAAAA	1463	
Db	44514	GAAGTAAAAAAAAAAAAAAAAAAAAA	44581	

RESULT	26
LOCUS	AC020910
DEFINITION	AC020910 Homo sapiens chromosome 19 clone CTF-2553L3, complete sequence.
VERSION	AC020910
KEYWORDS	AC020910.6 GI:15887301
SOURCE	Htg.
ORGANISM	human.
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 203250) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Unpublished 2 (bases 1 to 203250) DOE Joint Genome Institute. Direct Submission Submitted (12-JUN-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA 3 (bases 1 to 203250) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Submitted (28-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA 4 (bases 1 to 203250) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA On Oct 3, 2001 this sequence version replaced gi:1558584. Location/Qualifiers 1..203250



```
* 209494 209513: gap of unknown length
* 209514 210401: contig of 888 bp in length
* 210402 210421: gap of unknown length
* 210422 212117: contig of 1696 bp in length
* 212118 212137: gap of unknown length
* 212138 213793: contig of 1656 bp in length
* 213794 213813: gap of unknown length
* 213814 216128: contig of 2315 bp in length
* 216129 216148: gap of unknown length
* 216149 217683: contig of 1535 bp in length
* 217684 217703: gap of unknown length
* 217704 218899: contig of 1196 bp in length
* 218900 218919: gap of unknown length
* 218920 220247: contig of 1328 bp in length
* 220248 220267: gap of unknown length
* 220268 220706: contig of 439 bp in length
* 220707 220726: gap of unknown length
* 220727 220973: contig of 247 bp in length
* 220974 220993: gap of unknown length
* 220994 222103: contig of 1110 bp in length
* 222104 222123: gap of unknown length
* 222124 223614: contig of 1491 bp in length
* 223615 223634: gap of unknown length
* 223635 225741: contig of 2107 bp in length
* 225742 225761: gap of unknown length
* 225762 227521: contig of 1760 bp in length
* 227522 227541: gap of unknown length
* 227542 228671: contig of 1130 bp in length
* 228672 228691: gap of unknown length
* 228692 230160: contig of 1469 bp in length
* 230161 230180: gap of unknown length
* 230181 231658: contig of 1478 bp in length
* 231659 231678: gap of unknown length
* 231679 233124: contig of 1446 bp in length
* 233125 233144: gap of unknown length
* 233145 234325: contig of 1181 bp in length
* 234326 234345: gap of unknown length
* 234346 236158: contig of 1813 bp in length
* 236159 237006: contig of 828 bp in length
* 237007 237026: gap of unknown length
* 237027 238728: contig of 1702 bp in length.

FEATURES
source
1. 238728
/organism="Mus musculus"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-79P14"
/sex="male"
1. 31901
/note="assembly_name:Contig288"
misc_feature
31922..63057
/note="assembly_name:Contig287"
misc_feature
63078..83212
/note="assembly_name:Contig286"
misc_feature
83233..96794
/note="assembly_name:Contig285"
misc_feature
96815..111361
/note="assembly_name:Contig284"
misc_feature
111382..120966
/note="assembly_name:Contig283"
misc_feature
120987..129252
/note="assembly_name:Contig282"
misc_feature
129273..136700
/note="assembly_name:Contig281"
misc_feature
136721..143527
/note="assembly_name:Contig280"
misc_feature
143548..151277
/note="assembly_name:Contig279"
misc_feature
151298..157516
/note="assembly_name:Contig278"
misc_feature
157537..162065
```

```
misc_feature
/note="assembly_name:Contig277"
162086..165464
/note="assembly_name:Contig276"
clone_end:t7
vector_side:left"
misc_feature
165485..167866
/note="assembly_name:Contig275"
misc_feature
167887..170318
/note="assembly_name:Contig274"
misc_feature
170339..174287

Query Match
Best Local Similarity 100.0%; Pred. No. 5.1e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 19002 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 18965
```

```
RESULT 28
AX093307/c 396 bp DNA linear PAT 30-MAR-2001
LOCUS
DEFINITION Sequence 125 from Patent WO0118046.
ACCESSION AX093307
VERSION AX093307.1 GI:13509756
KEYWORDS
SOURCE
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

```
REFERENCE
1 (bases 1 to 396)
AUTHORS Xu,J. and Stolk,J.A.
TITLE Ovarian tumor sequences and methods of use therefor
JOURNAL Patent: WO 0118046-A 125 15-MAR-2001;
CORIXA CORPORATION (US)
```

Location/Qualifiers

source

1. 396 /organism="Homo sapiens"

misc\_feature /db\_xref="taxon:9606"

1. 396 /note="n = A,T,C or G"

BASE COUNT 115 a 71 c 60 g 125 t 25 others

ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 396;

Best Local Similarity 100.0%; Pred. No. 1.4e-09;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 40 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4
```

RESULT 29

BC009480

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

BC009480 445 bp mRNA linear PRI 12-JUL-2001  
Homo sapiens, SEC61, gamma subunit (S. cerevisiae), clone MGC:15510  
IMAGE:3954682, mRNA, complete cds.  
BC009480  
BC009480.1 GI:14550464  
MGC.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Strausberg,R.  
Direct Submission  
Submitted (25-JUN-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA



REMARK  
COMMENT  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: ATCC/DC/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
Contact: [amadan@systemsbiology.org](mailto:amadan@systemsbiology.org)  
Anup Madan, Rachel Dickhoff, Jessica Fahy, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 25 Row: j Column: 2  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 4092055.

FEATURES  
source  
Location/Qualifiers  
1..445  
/organism="Homo sapiens"  
/db\_xref="LocustID:20335"  
/db\_xref="taxon:9606"  
/clone="MGC:15310 IMAGE:3954682"  
/tissue\_type="Skin, melanotic melanoma."  
/clone\_1lb="NIH\_MGC\_20"  
/lab\_host="DH10B-R"  
/note="Vector: POTB7"  
43..249  
/codon\_start=1  
/product="SEC61, gamma subunit (S. cerevisiae)"  
/protein\_id="AAH0480.1"  
/db\_xref="GI:14550465"  
/translation="MDQVMOFVEPSRFVKSIRLVKRCIKRDKKEFGKIMATAIGF  
ALMGFIGFVKLIHIPINNIIYGC"  
56 c 97 g 137 t

CDS  
BASE COUNT 155 a 56 c 97 g 137 t  
ORIGIN

Query Match 2.5%; Score 37; DB 9; Length 445;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAA  
DB 400 AAGTAAAAA  
1463  
436

RESULT 30  
LOCUS E01762 508 bp RNA linear PAT 29-SEP-1997  
DEFINITION Variant aequorin gene.  
ACCESSION E01762  
VERSION E01762.1 GI:2170015  
KEYWORDS JP 1988291586-A/1.  
SOURCE unclassified.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 508)  
AUTHORS Inoue,S., Kurose,K. and Sakaki,Y.  
TITLE VARIANT Aequorin GENE  
JOURNAL Patent: JP 1988291586-A 1 29-NOV-1988;  
CHISSO CORP  
OS Medusa  
PN JP 1988291586-A/1  
PD 29-NOV-1988  
PF 23-MAY-1987 JP 1987126374  
PI INOUE SATOSHI, KUROSE KOICHI, SAKAKI YOSHIYUKI PC  
C12N15/00//C07K13/00,C12P21/02,(C12P21/02,C12R1.19); CC  
strandedness: Double;  
CC topology: linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: clone=PAQ440;  
FH Key Location/Qualifiers

FH  
FT misc\_feature 1..958  
FT FT  
FT variation /note='aequorin gene'  
FT replace(659,'c') /note='variant aequorin gene' FT variation  
FT replace(590,'c') /note='variant aequorin gene' FT variation  
FT replace(674,'c') /note='variant aequorin gene' FT variation  
FT replace(569,'c') /note='variant aequorin gene' FT variation

FEATURES  
source  
Location/Qualifiers  
1..508  
/organism="unidentified"  
/db\_xref="taxon:32644"  
191 a 75 c 133 g 109 t

BASE COUNT 191 a 75 c 133 g 109 t  
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 508;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAA  
DB 471 AAGTAAAAA  
1463  
507

RESULT 31  
LOCUS E01765 508 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA sequence of equine.  
ACCESSION E01765  
VERSION E01765.1 GI:2170018  
KEYWORDS JP 1988291593-A/1.  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 508)  
AUTHORS Inoue,S., Kurose,K. and Sakaki,Y.  
TITLE PRODUCTION OF VARIANT Aequorin PROTEIN  
JOURNAL Patent: JP 1988291593-A 1 29-NOV-1988;  
CHISSO CORP  
OS Artificial gene  
OC Artificial sequence; Genes.  
PN JP 1988291593-A/1  
PD 29-NOV-1988  
PF 23-MAY-1987 JP 1987126373  
PI INOUE SATOSHI, KUROSE KOICHI, SAKAKI YOSHIYUKI PC  
C12P21/02//C07K13/00,C12N15/00,(C12P21/02,C12R1.19); CC  
strandedness: Double;  
CC topology: linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH key Location/Qualifiers  
FH FT  
FH CDS 1..958  
FT FT  
FT /product='equine'.  
FEATURES  
source  
Location/Qualifiers  
1..508  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
191 a 75 c 133 g 109 t

```

RESULT 32
E08516 LOCUS E08516 826 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA encoding style-specific S-ribonuclease.
ACCESSION E08516
VERSION E08516.1 GI:2176631
KEYWORDS JP 1994335389-A/1.
SOURCE Lycopersicon peruvianum.
ORGANISM Lycopersicon peruvianum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 826)
Tei,I., Nakada,K., Ito,T., Horiuchi,H., Ota,A., Takagi,M.,
Tsubura,H., Tanaka,H. and Ishiguro,T.,
S-RIBONUCLEASE SPECIFIC TO STYLE AND DNA SEQUENCE CODING THEREFOR
Patent: JP 1994335389-A 1 06-DEC-1994;
KAGOME CO LTD
COMMENT OS Lycopersicon peruvianum (tomato)
PN JP 1994335389-A/1
PD 06-DEC-1994
PE 27-MAY-1993 JP 1993126286
PI TEI ITSUKIYON, NAKADA KENGO, ITO TORU, HORIUCHI HIROYUKI, PI
OTA AKIMORI,
PI TAKAGI MASAMICHI, TSUBURA HIROKAZU, TANAKA HIROSHI, PI
ISHIGURO YUKIO
PC C12N9/22,C12N15/52;
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..826
FT FT /organism='Lycopersicon peruvianum' FT
FT FT /tissue-type='pistil'
FT FT /product='style-specific S-ribonuclease'.
FEATURES
source 1..826
location/Qualifiers
1..826
/organism="Lycopersicon peruvianum"
/db_xref="taxon:4082"
BASE COUNT 309 a 150 c 131 g 236 t
ORIGIN
Query Match 2.5%; Score 37; DB 6; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 774 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810

RESULT 33
E12586 LOCUS E12586 826 bp DNA linear PAT 24-JUN-1998
DEFINITION cDNA encoding S-ribonuclease.
ACCESSION E12586
VERSION E12586.1 GI:3251418
KEYWORDS JP 1997028381-A/3.
SOURCE Lycopersicon peruvianum.
ORGANISM Lycopersicon peruvianum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 826)
Tei,I., Minami,K. and Takagi,M.,
S-RIBONUCLEASE GENE AND PROMOTER SEQUENCE
Patent: JP 1997028381-A 3 04-FEB-1997;
TEI ITSUKIYON, MINAMI KOUICHI, TAKAGI MASAMICHI
OS Lycopersicon peruvianum
PN JP 1997028381-A/3

```

```

PD 04-FEB-1997
PE 24-JUL-1995 JP 1995187557
PI TEI ITSUKIYON, MINAMI KOUICHI, TAKAGI MASAMICHI PC
C12N15/09,C07H21/04,C12N1/21//A01H1/00,C12N5/10,C12N9/22, PC
(C12N1/21,
PC C12R1.19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..826
FT FT /organism='Lycopersicon peruvianum' FT CDS
FT FT 1..489
FT FT /product='S-ribonuclease'
FT FT 3'UTR
FT FT Location/Qualifiers
1..826
/organism="Lycopersicon peruvianum"
/db_xref="taxon:4082"
BASE COUNT 309 a 150 c 131 g 236 t
ORIGIN
Query Match 2.5%; Score 37; DB 6; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 774 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810

RESULT 34
S65048 LOCUS S65048 826 bp mRNA linear PLN 29-OCT-1993
DEFINITION self-incompatibility associated S-glycoprotein [clone II]
[Lycopersicon peruvianum=tomatoes, PI 126441, style, Slla-plant,
mRNA Partial, 826 nt].
ACCESSION S65048
VERSION S65048.1 GI:410002
KEYWORDS
SOURCE Lycopersicon peruvianum
ORGANISM Lycopersicon peruvianum
Peruvian tomato style PI 126441 Slla-plant.
Lycopersicon peruvianum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 826)
Chung,I.K., Nakada,K., Tanaka,H., Ito,T., Horiuchi,H., Ohta,A. and
Takagi,M.,
Identification of cDNA clones coding for the style specific
Slla-glycoprotein gene associated with gametophytic
self-incompatibility in tomato (Lycopersicon peruvianum)
Biosci. Biotechnol. Biochem. 57 (7), 1172-1176 (1993)
93372490
GenBank staff at the National Library of Medicine created this
entry [NCBI g1bbsq 136760] from the original journal article.
This sequence comes from Fig. 1.
FEATURES
source 1..826
location/Qualifiers
1..826
/organism="Lycopersicon peruvianum"
/db_xref="taxon:4082"
1..826
/partial
1
/gene="self-incompatibility associated S-glycoprotein"
BASE COUNT 309 a 152 c 129 g 236 t
ORIGIN
Query Match 2.5%; Score 37; DB 8; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 ||||||||||||||||||||||||||||||||||||  
 DB 774 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810

RESULT 35  
 AF064030 829 bp mRNA linear PLN 22-MAR-1999  
 LOCUS Helianthus tuberosus lectin 2 mRNA, complete cds.  
 ACCESSION AF064030  
 VERSION AF064030.1 GI:4091896  
 KEYWORDS  
 SOURCE Helianthus tuberosus.  
 ORGANISM Helianthus tuberosus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; easterids II; Asterales; Asteraceae; Asteroideae;  
 Heliantheae; Helianthus.  
 1 (bases 1 to 829)  
 Van Damme, E.-J., Batte, A., Mazard, A.M., Verhaert, P., Horman, A.,  
 Degray, H., Rouge, P. and Peumans, W.J.  
 Characterization and molecular cloning of the lectin from  
 Helianthus tuberosus  
 Eur. J. Biochem. 259 (1-2), 135-142 (1999)  
 99115433  
 2 (bases 1 to 829)  
 Van Damme, E.-J.M. and Peumans, W.J.  
 Direct Submission  
 Submitted (08-MAY-1998) Laboratory of Phytopathology and Plant  
 Protection, Catholic University of Leuven, Willem de Croylaan 42,  
 Leuven Herestraat, Brabant 3001, Belgium  
 Location/Qualifiers  
 source 1. 829  
 /organism="Helianthus tuberosus"  
 /db\_xref="taxon:4233"  
 82. .525  
 /note="LECHeltuba2; agglutinin; helituba"  
 /codon\_start=1  
 /product="lectin 2"  
 /protein\_id="AA011578.1"  
 /db\_xref="GI:4091897"  
 /translation="MAASDTCVQAGPGCGGGRWLOTAGHKITSIIITKGCIFSI  
 OFVTKDNIETVSGGQGVGQDKAETITTEADDEDITATISGFAYVQMTVVSITFTOT  
 NKVYGPFGVAGSRFSLPLTKGFAGFENSDVLDISIGVVP"

BASE COUNT 258 a 150 c 188 g 233 t

ORIGIN

Query Match 2.5%; Score 37; DB 8; Length 829;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 ||||||||||||||||||||||||||||||||||||  
 DB 785 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 821

RESULT 36  
 ATP27A 835 bp mRNA linear PLN 14-MAR-1997  
 LOCUS A. thaliana mRNA for peroxidase ATP27a, partial.  
 ACCESSION Y11792  
 VERSION Y11792.1 GI:1890318  
 KEYWORDS peroxidase.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 835)  
 Weidner, K.G., Jespersen, H.M., Kjaerregaard, I.V.M., Justesen, A.F.,  
 Oestergaard, L., Abelskov, A.K., Jensen, R.B., Hansen, L.N. and  
 Rasmussen, S.K.  
 From expressed sequence tags to structure, function, evolution and

JOURNAL unpublished  
 REFERENCE 2 (bases 1 to 835)  
 AUTHORS Justesen, A.F.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-MAR-1997) A.F. Justesen, University of Copenhagen,  
 Dept. of Protein Chemistry, Oester Farimagsgade 2a, DK-1353  
 Copenhagen K, DENMARK  
 Location/Qualifiers  
 source 1. 835  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 /clone="2011187"  
 /cell\_line="Columbia"  
 /clone\_lib="lambda PR12"  
 /dev\_stage="roots"  
 /dev\_stage="stems"  
 /dev\_stage="rosettes"  
 /dev\_stage="flowers"  
 /dev\_stage="etiolated seedlings"  
 <1. .616  
 /EC\_number="1.11.1.7"  
 /codon\_start=2  
 /product="peroxidase ATP27a"  
 /protein\_id="CAA7248.1"  
 /db\_xref="GI:1890319"  
 /db\_xref="SPTREMBL:p93727"  
 /translation="SCADVATAIAARVYVTLSCGPVWSVLGGRKDGITISANETRLNP  
 PTEVNSQLISFPAARGLSVKDMVYVLTSGHTIGFSHCSSESLQNSKPHDIDPMNY  
 AAFQTLKKCPRTSNRGNAGVYLDSTSFVDMVYKKQLSGVGSQALDGSRT  
 KWIETPDAQOKAFPREFAASVYKLNFGVKETGQVYVNTREVN"

BASE COUNT 285 a 172 c 166 g 212 t

ORIGIN

Query Match 2.5%; Score 37; DB 8; Length 835;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 ||||||||||||||||||||||||||||||||||||  
 DB 768 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 804

RESULT 37  
 BC018351 848 bp mRNA linear ROD 06-DEC-2001  
 LOCUS Mus musculus, proteasome (prosome, macropain) subunit, beta type 1,  
 clone MGC:5916 IMAGE:3584538, mRNA, complete cds.  
 ACCESSION BC018351  
 VERSION BC018351.1 GI:17390819  
 KEYWORDS MGC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 848)  
 Strausberg, R.  
 Direct Submission  
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: c9apbs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/

Contact: villalonebcm.tmc.edu.  
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,  
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the T.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRK Plate: 10 Row: 1 Column: 19  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA g1: 7242196.

## FEATURES

Location/Qualifiers  
1..848  
/organism="Mus musculus"  
/db\_xref="LocusID:19170"  
/db\_xref="taxon:10090"  
/clone="MGC:5916 IMAGE:3584538"  
/tissue\_type="Mammary tumor. Metallothionien-TGF alpha  
model. 10 month old virgin mouse. Taken by biopsy."  
/clone\_lib="NCI CGAP\_Mam1"  
/lab\_host="DH10B"  
/note="vector: PCMV-SPORT6"  
16..738  
/codon\_start=1  
/product="proteasome (prosome, macropain) subunit, beta  
type 1"  
/protein\_id="AAH18351.1"  
/db\_xref="GI:17390820"  
/translation="MSTVAAYRDERELGMPHSGAGVQLRFSPYANGTVLAIAG  
EDFSIVASDTRLSEGSIFHTRDSPCYRLTDVTYIGSGFGDCITLTKIIEARLKM  
KHSNNAMTGTIAAMLSTILSRFPFYVNIIGLDESGKAVSFEDVGSYORD  
SPKAGSASAMLOPLDNOVGFKMNQNEHVPLTIDRAMRLVKDFISAARDVYTG  
ALRICVMEGRREFTVPLRKD"

## CDS

BASE COUNT 259 a 179 c 214 g 196 t  
ORIGIN

Query Match 2.5%; Score 37; DB 10; Length 848;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
Db 795 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 831

RESULT 38  
AEVQA440X 925 bp mRNA linear INV 18-MAR-1994  
LOCUS Aequorea victoria aequorin (AQ440) mRNA, complete cds.  
ACCESSION 129571  
VERSION 129571.1 GI:461374  
KEYWORDS aequorin; luminescent protein; oxygenase; photoprotein.  
SOURCE Aequorea victoria cDNA to mRNA.  
ORGANISM Aequorea victoria  
Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;  
Aequoreidae; Aequorea.  
REFERENCE 1 (bases 1 to 925)  
AUTHORS Inouye, S., Noguchi, M., Sakaki, Y., Takagi, Y., Miyata, T., Iwanaga, S.  
and Tsuji, F. I  
TITLE Cloning and sequence analysis of cDNA for the luminescent protein  
aequorin  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 82, 3154-3158 (1985)  
MEDLINE 85216460  
FEATURES  
location/Qualifiers  
1..925  
/organism="Aequorea victoria"  
/db\_xref="taxon:6100"  
/cell\_type="photocyte"  
1..925  
/gene="AQ440"  
82..672  
/gene="AQ440"  
/note="EF-hand/Calcium binding site amino acids 31..42  
EF-hand/Calcium binding site amino acids 124..133

EF-hand/Calcium binding site amino acids 160..171"

/codon\_start=1  
/product="aequorin"  
/protein\_id="AA27720.1"  
/db\_xref="GI:461375"  
/translation="MRSKQYSVKLTSPDNPRTGRHKHMFELDVNHNKISLDEMY  
YKASDIVINNLGATPEOAKRHRDAVEFGAGMYCVETDPAYTEGKRLATPELE  
KVAKNEPTLRITWGDAFDIVKDDONGATLDEMKAYTAAAGIIQSECEETFRVCD  
IDSGOLDVDENITROHLGFMVMTDPAECERLYGAVP"

BASE COUNT 362 a 156 c 178 g 229 t  
ORIGIN

Query Match 2.5%; Score 37; DB 3; Length 925;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
Db 858 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 894

RESULT 39  
AR108851 958 bp DNA linear PAT 14-FEB-2001  
LOCUS AR108851  
DEFINITION Sequence 5 from patent US 6113886.  
ACCESSION AR108851  
VERSION AR108851.1 GI:12825127  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 958)  
AUTHORS Bryan, B.  
TITLE Bioluminescent novelty items  
JOURNAL Patent: US 6113886-A 5 05-SEP-2000;  
FEATURES  
location/Qualifiers  
1..958  
/organism="unknown"

BASE COUNT 363 a 156 c 210 g 229 t  
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 958;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 40  
AR119320 958 bp DNA linear PAT 16-MAY-2001  
LOCUS AR119320  
DEFINITION Sequence 5 from patent US 6152358.  
ACCESSION AR119320  
VERSION AR119320.1 GI:14102019  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 958)  
AUTHORS Bryan, B.  
TITLE Bioluminescent novelty items  
JOURNAL Patent: US 6152358-A 5 28-NOV-2000;  
FEATURES  
location/Qualifiers  
1..958  
/organism="unknown"

BASE COUNT 363 a 156 c 210 g 229 t  
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 958;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 41  
AR151721 958 bp DNA linear PAT 08-AUG-2001  
LOCUS Sequence 5 from patent US 6232107.  
DEFINITION AR151721  
ACCESSION AR151721  
VERSION AR151721.1 GI:15117771  
KEYWORDS  
SOURCE Unknown.  
ORGANISM  
REFERENCE  
1 (bases 1 to 958)  
Bryan,B.J. and Szent-Gyorgyi,C.  
Luciferases, fluorescent proteins, nucleic acids encoding the  
luciferases and fluorescent proteins and the use thereof in  
diagnostics, high throughput screening and novelty items  
Patent: US 6232107-A 5 15-MAY-2001;  
Location/Qualifiers  
1..958  
/organism="unknown"  
BASE COUNT 363 a 156 c 210 g 229 t  
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 958;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 42  
AX250563 958 bp DNA linear PAT 05-OCT-2001  
LOCUS Sequence 5 from Patent WO0168824.  
DEFINITION AX250563  
ACCESSION AX250563  
VERSION AX250563.1 GI:15984303  
KEYWORDS  
SOURCE Aequorea victoria.  
ORGANISM Aequorea victoria  
Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;  
Aequoreidae; Aequorea.  
1 (bases 1 to 958)  
Bryan,B.J., Szent-Gyorgyi,C. and Szczepaniak,W.  
Renilla reniformis fluorescent proteins, nucleic acids encoding the  
fluorescent proteins and the use thereof in diagnostics, high  
throughput screening and novelty items  
Patent: WO 0168824-A 5 20-SEP-2001;  
Protein, Ltd. (US); Bryan, Bruce J. (US)  
Location/Qualifiers  
1..958  
/organism="Aequorea victoria"  
/db\_xref="taxon:6100"  
115..705  
/note="Apoequorin-encoding gene"  
/codon\_start=1  
/protein\_id="CAC93773.1"  
/db\_xref="GI:15984303"  
translation="MTSKQSYVKLISDFDNFRWIGRHKHMFNLDVNHNGKISIDEMV  
YKASDIYNNIGATPPEQAKRKHDAVEAFGGAGMKYGVETDMPATIEGKKIATDELE  
KYAKKEPLIRIMGDALEFDYDKDNGAIIIDEMWAAYKAGIIGSSSDCETFEVCD  
IDESGQGLVDMDTRHIGLFWYIMDPACEKLYGGAVP"  
BASE COUNT 363 a 156 c 210 g 229 t  
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 958;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 43  
E01573 958 bp RNA linear PAT 29-SEP-1997  
LOCUS CDNA sequence of wild type of aquolin gene PAQ440.  
DEFINITION E01573  
ACCESSION E01573  
VERSION E01573.1 GI:2169826  
KEYWORDS JP 1986098387-A/1.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 958)  
Inoue,S., Kurose,K. and Sakaki,Y.  
MUTANT AEQUORIN GENE  
Patent: JP 1986098387-A 1 28-APR-1988;  
CHISSO CORP  
PN JP 1988098387-A/1  
PD 28-APR-1988  
PI INOUE SATOSHI, KUROSE KOICHI, SAKAKI YOSHIYUKI PC  
C12N15/00//C12P21/00,G01N21/76;  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH Key Location/Qualifiers  
FT misc.feature 1..958  
FT /gene="wild type of aquolin gene PAQ440".  
Location/Qualifiers  
1..958  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 363 a 156 c 210 g 229 t  
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 958;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 44  
E01575 958 bp RNA linear PAT 29-SEP-1997  
LOCUS CDNA sequence of wild type of aquolin gene PAQ440.  
DEFINITION E01575  
ACCESSION E01575  
VERSION E01575.1 GI:2169828  
KEYWORDS JP 1986098398-A/1.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 958)  
Inoue,S., Kurose,K. and Sakaki,Y.  
PRODUCTION OF MUTANT AEQUORIN PROTEIN  
Patent: JP 1986098398-A 1 28-APR-1988;  
CHISSO CORP  
PN JP 1988098398-A/1  
PD 28-APR-1988  
PI INOUE SATOSHI, KUROSE KOICHI, SAKAKI YOSHIYUKI PC  
C12P21/02,C12N15/00//C07K13/00,(C12P21/02,C12R1:19); CC

strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH key Location/Qualifiers  
FH misc-feature 1..958  
FT /gene='wild type of aquolin gene PAQ440'.  
Location/Qualifiers  
1..958  
source /organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 363 a 156 c 210 g 229 t  
ORIGIN

Query Match 2.58; Score 37; DB 6; Length 958;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 45  
E02319 958 bp RNA linear PAT 29-SEP-1997  
LOCUS DNA encoding luminescent protein aqualine.  
DEFINITION E02319  
ACCESSION E02319.1 GI:2170555  
VERSION JP 1990096597-A/1.  
KEYWORDS Aequorea victoria.  
SOURCE Aequorea victoria.  
ORGANISM Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;  
Aequoreidae; Aequorea.  
REFERENCE 1 (bases 1 to 958)  
AUTHORS Takagi,Y., Sakaki,Y., Inoue,S., Noguchi,M., Iwanaga,S., Miyata,T.  
and Furederitsuku,A.T.  
TITLE BIOSYNTHESIS OF PHOTOPROTEIN AEQUORIN  
JOURNAL Patent: JP 1990096597-A 1 09-Apr-1990;  
CHISSO CORP  
COMMENT OS Aequorea victoria  
PN JP 1990096597-A/1  
PD 09-APR-1990  
PF 08-NOV-1988 JP 1989113000  
PI TAKAGI YASUTAKA, SAKAKI YOSHIYUKI, INOUE SATOSHI, PI NOGUCHI MASATO,  
PI IWANAGA SADAKI, MIYATA TOSHIYUKI, FUREDERITSUKU AI TSUJI PC  
C07K13/00,C12P21/02//C12N15/12,(C12P21/02,C12R1:19); CC  
strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: tissue\_type=plaus;  
CC \*source: clone=paQ440;  
CC Feature is identified by experimental;  
FH Key Location/Qualifiers  
FH 5'UTR 1..114  
FT CDS 115..705  
FT /product='luminescent protein aqualine' FT  
FT mat\_peptide 115..702  
FT /product='luminescent protein aqualine' FT  
3'UTR 706..958  
Location/Qualifiers  
1..958  
source /organism="Aequorea victoria"  
/db\_xref="taxon:6100"  
BASE COUNT 362 a 155 c 212 g 229 t  
ORIGIN

Query Match 2.58; Score 37; DB 6; Length 958;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

Search completed: November 5, 2002, 15:59:48  
Job time : 4291.32 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 10:52:52; Search time 174.666 Seconds

(without alignments)  
14380.822 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463

Sequence: 1 caccagaatagctcgcgcgc.....aaaaaaaaaaaaaaaa 1463

Scoring table:

OLIGO-NUC  
Gapop 60.0, Gapext 60.0

Searched: 1736436 segs, 858457221 residues

Word size: 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: N.GeneSeq\_032802.\*

1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1463	100.0	1463	21	Maize Rad2/FEN-1 c
2	1360	93.0	1541	21	Maize Rad2/FEN-1 c
3	1267	86.6	1381	21	Maize Rad2/FEN-1 c
4	894	61.1	1478	21	Maize Rad2/FEN-1 c
5	40	2.7	584	21	Human prostate can
6	39	2.7	288	22	Human cervical can
7	39	2.7	305	22	Human cervical can
8	38	2.6	691	20	Human prostate tum
9	38	2.6	2381	21	Lung cancer associ

10	38	2.6	8711	22	AA526809	Human genomic DNA
11	37	2.5	302	21	AA598700	Human colon cancer
12	37	2.5	342	22	AA183214	Human polynucleoti
13	37	2.5	396	22	AA594934	Human ovarian canc
14	37	2.5	411	22	AA188595	Human polynucleoti
15	37	2.5	412	22	AA183922	Human polynucleoti
16	37	2.5	426	22	AA119865	Human breast cance
17	37	2.5	524	21	AA596571	Noncoding region o
18	37	2.5	699	22	AA5964213	Human secreted pro
19	37	2.5	706	21	AA596563	Human secreted pro
20	37	2.5	826	16	AA5979736	Flower style-speci
21	37	2.5	826	18	AA5964553	Tomato S-ribonucle
22	37	2.5	957	11	AA5964431	Aequorin gene, Ae
23	37	2.5	958	22	AA5962233	Aequorin-encodi
24	37	2.5	958	24	AA5962218	Aequorea victoria
25	37	2.5	959	9	AA5961534	PAQ440 aequorin ge
26	37	2.5	1093	21	AA597825	Human cancer assoc
27	37	2.5	1159	22	AA592547	Human secreted pro
28	37	2.5	1252	22	AA5929087	CDNA encoding for
29	37	2.5	1293	24	AA5963048	Cell death protect
30	37	2.5	1362	19	AA5932421	Homo sapiens clone
31	37	2.5	1362	22	AA598437	Human CDNA clone C
32	37	2.5	1380	22	AA593111	Human secreted pro
33	37	2.5	1629	21	AA59580	Human secreted pro
34	37	2.5	1721	22	AA5963820	Human secreted pro
35	37	2.5	1800	22	AA5907579	Human secreted pro
36	37	2.5	2108	21	AA596570	A core 2 beta-1,6-
37	37	2.5	2198	21	AA593424	CDNA encoding huma
38	37	2.5	2260	22	AA591914	Human neuroblastom
39	37	2.5	2260	22	AA598068	Human neuroblastom
40	37	2.5	2297	21	AA593309	Sequence encoding
41	37	2.5	2602	21	AA512416	CDNA encoding a hu
42	37	2.5	2827	22	AA596812	Human LP8, a PDGF-
43	37	2.5	2836	19	AA590744	C. felis esterase,
44	37	2.5	2836	19	AA590745	C. felis esterase,
45	37	2.5	2836	19	AA590745	Ctenocephalides fe

# ALIGNMENTS

RESULT 1	AAA27923	standard; CDNA: 1463 BP.
ID	AAA27923	
AC	AAA27923	
XX		
DT	12-SEP-2000	(first entry)
XX		
DE	Maize Rad2/FEN-1 cDNA.	
XX		
KW	Maize, Rad2/FEN-1; transgenic plant; male sterile plant;	
KW	endonuclease; exonuclease; DNA repair; gene targeting; ss.	
XX		
OS	Zea mays.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	85..1224
XX		/*tag= a
XX		
PN	WO200036109-A1.	
PD	22-JUN-2000.	
XX		
PF	16-NOV-1999;	99WO-US27147.
XX		
PR	15-DEC-1998;	98US-0112332.
XX		
PA	(PION-) PIONEER HI-BRED INT. INC.	
XX		
PI	Mahajan PB;	
XX		
DR	WPI; 2000-452026/39.	

DR P-PSDB; AAY95307.

aa Maize RAB2/FEN-1 nucleic acids and proteins useful for modulating DNA  
 pt recombination and repair in transgenic plants, e.g. for gene targeting  
 pt and the production of male sterile plants -  
 xx  
 ps Claim 1; Page 69-71; 85pp: English.

Claim 1; Page 69-71; 85pp; English.

The present sequence is that of maize cDNA coding for RAD2/FEN-1 (see AY93307). The corresponding RNA was isolated from immature ear tissue from 2 ears of a B73 maize line. Rad2/FEN-1 is a structure specific endonuclease which under certain conditions also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to produce the Rad2/FEN-1 polypeptides in transgenic plant cells. The protein is involved in the regulation of DNA repair and recombination in plant systems and therefore may be used for improving gene targeting during further recombinant DNA protocols involving plants. RAD2/FEN-1 endonucleolytic activity is essential in DNA replication and nucleotide excision and repair reactions. The exolytic activity is involved in double strand break repair and end joining. The protein is also useful in strand exchange reactions during homologous recombination. These functions may be useful in gene targeting and in the production of male sterile plants. The efficacy of gene targeting can be improved by the overexpression of exogenous Rad2/FEN-1 while male sterile plants can be produced by the down regulation of Rad2/FEN-1 expression.

Sequence 1463 BP; 466 A; 292 C; 361 G; 344 T; 0 other;

Query Match	100.0%;	Score 1463;	DB 21;	Length 1463;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1463: Conservative	0;	Mismatches	0;	Gaps 0;

Oy	1	CACGAGAAATAGTGTGGGGGGCGGGGTTTCTTGCGGCATCTCCGAGCTCAGACGGCGCCGGCA	60
Db	1	CACGAGAAATAGTGTGGGGGGCGGGGTTTCTTGCGGCATCTCCGAGCTCAGACGGCGCCGGCA	60
Oy	61	CCCGGCACAGCCCGCCGAGAGAGATGGGATCAAGGGTTTGACGAACCTGCTGGCGAC	120
Db	61	CCCGGCACAGCCCGCCGAGAGAGATGGGATCAAGGGTTTGACGAACCTGCTGGCGAC	120
Oy	121	AATGGCCCCAAGGCGATTAAGGAGCAGAGATTGCAAGACTACTTCGGCCGCAAAATCGGC	180
Db	121	AATGGCCCCAAGGCGATTAAGGAGCAGAGATTGCAAGACTACTTCGGCCGCAAAATCGGC	180
Oy	181	GTGCGAGCCGAGTATGAGCATATACAGTACTCGATTGATTTGAGTGAAGAGCAGCATGAA	240
Db	181	GTGCGAGCCGAGTATGAGCATATACAGTACTCGATTGATTTGAGTGAAGAGCAGCATGAA	240
Oy	241	ACTCTTCACAAATGAAGAGCTGTGTGAAGTCACTAGTATTTGCAAGAGATGTTCAACCGACA	300
Db	241	ACTCTTCACAAATGAAGAGCTGTGTGAAGTCACTAGTATTTGCAAGAGATGTTCAACCGACA	300
Oy	301	ATAAGATTACTGGAAAGCGGGAATCAAGCCAGTTTATGTTTTGATGSCAAGCCTCCTGAT	360
Db	301	ATAAGATTACTGGAAAGCGGGAATCAAGCCAGTTTATGTTTTGATGSCAAGCCTCCTGAT	360
Oy	361	ATGAAGAACAAGAGGCTGTCTTAAAGATATCTCAAAAACAGATGATGCAACCAAGATCTG	420
Db	361	ATGAAGAACAAGAGGCTGTCTTAAAGATATCTCAAAAACAGATGATGCAACCAAGATCTG	420
Oy	421	ACTGAGGCAGTAGAGGTAGAGATTAAGATGCGATTGAAAAATTGAGCAAGAGACTGTA	480
Db	421	ACTGAGGCAGTAGAGGTAGAGATTAAGATGCGATTGAAAAATTGAGCAAGAGACTGTA	480
Oy	481	AAGGTCAACAAGGCACACACACGAGATTGTAAAGGGCTATTAAAGCTTATGCGGGTTCTT	540
Db	481	AAGGTCAACAAGGCACACACACGAGATTGTAAAGGGCTATTAAAGCTTATGCGGGTTCTT	540
Oy	541	GTTGTAGAGGCACCTTCTGAGCAGAGAGAGAAATGCGAGCCCTTGGATTAAGCATAG	600
Db	541	GTTGTAGAGGCACCTTCTGAGCAGAGAGAGAAATGCGAGCCCTTGGATTAAGCATAG	600

Oy	601	GATGTCCTGTCCCTCAGACATATGACTGCCCTTACTTTTGSGGSCCTCAAGGGTCCCTT	660
Db	601	GTGTTCCTGTCTGCCTCAGAAGATAAGAACTCCCTTACTTTTGSGGCTCCACAGGTTCTTT	660
Oy	661	CGCATTTAATGGATCGCAACTTCGAAGAAATACCCTGTGATGSAATTGGATGTGCCAAG	720
Db	661	CCTCATTTTAATGATGCCAAGTTCCAAAGAAAATACCTCGATGATGGAATTGATGTTGCCAAG	720
Oy	721	GTMTTGAAGAGCTTGAATCTCACCATGAGCCAGCTTCATTGATTTTGTGATCCCTGTGGA	780
Db	721	GTMTTGAAGAGCTTGAATCTCACCATGAGCCAGCTTCATTGATTTTGTGATCCCTGTGGA	780
Oy	781	TGTGACTATTGTGATPACATCAAGAGGTATTCGGGGGGCAAAACAGCTCTGMAACTTATTCGT	840
Db	781	TGTGACTATTGTGATPACATCAAGAGGTATTCGGGGGGCAAAACAGCTCTGMAACTTATTCGT	840
Oy	841	CACATAGGGCTCATAGAAAGACATCTTGGAGAACTCTTAATAAAGACAGATATCAAAATTCCT	900
Db	841	CACATAGGGCTCATAGAAAGACATCTTGGAGAACTCTTAATAAAGACAGATATCAAAATTCCT	900
Oy	901	GAGGACTGGCTTCACCAAGAAAGCTCCGAGCTTTTCAAGAGACCTTAATGTACAATTGGAT	960
Db	901	GAGGACTGGCTTCACCAAGAAAGCTCCGAGCTTTTCAAGAGACCTTAATGTACAATTGGAT	960
Oy	961	ATTCTCTGAGCTAAATGGACTGCACCTGATGAGAGGAGTCTATAAGTTTTCTGTGTAATA	1020
Db	961	ATTCTCTGAGCTAAATGGACTGCACCTGATGAGAGGAGTCTATAAGTTTTCTGTGTAATA	1020
Oy	1021	GATTAATGCTTTCACGAAAGATCGGGGTACAAAGGCCATAGAGAAAGTCAAAATCTGCCAAG	1080
Db	1021	GATTAATGCTTTCACGAAAGATCGGGGTACAAAGGCCATAGAGAAAGTCAAAATCTGCCAAG	1080
Oy	1081	AATTAATCGTGCACAAAGAAAGCTGTAGTCCTTTTTCAAGCCAACTGCCCACCATCAGCA	1140
Db	1081	AATTAATCGTGCACAAAGAAAGCTGTAGTCCTTTTTCAAGCCAACTGCCCACCATCAGCA	1140
Oy	1141	CCGCTTAAACGAGAGAGACTTCCGATTAANAACAAGGCAAGCTGCGAACCAAGAAACA	1200
Db	1141	CCGCTTAAACGAGAGAGACTTCCGATTAANAACAAGGCAAGCTGCGAACCAAGAAACA	1200
Oy	1201	AAGGCTGGGGAAGAAAGAAATAATCTTGGATGCTGATGTACAACTAGACTACGAAG	1260
Db	1201	AAGGCTGGGGAAGAAAGAAATAATCTTGGATGCTGATGTACAACTAGACTACGAAG	1260
Oy	1261	CACGGTGGCGTGATCACTTGCCTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTT	1320
Db	1261	CACGGTGGCGTGATCACTTGCCTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTT	1320
Oy	1321	GGTAAGAAGTTTCTCATGTGTTTCAAGCTGGGGTAGTAGTGTGTTGAAGAGATTGCTG	1380
Db	1321	GGTAAGAAGTTTCTCATGTGTTTCAAGCTGGGGTAGTAGTGTGTTGAAGAGATTGCTG	1380
Oy	1381	TACCAAGTAACAAACTTATCGCTGTTTTTTTACTCTTGTGCTTTGNAGTAAAAAAA	1440
Db	1381	TACCAAGTAACAAACTTATCGCTGTTTTTTTACTCTTGTGCTTTGNAGTAAAAAAA	1440
Oy	1441	AAAAAAAAAAAAAAAAAAAAAAAA 1463	
Db	1441	AAAAAAAAAAAAAAAAAAAAAAAA 1463	
<hr/>			
RESULT 2			
AAA27924			
ID	AAA27924	standard; cDNA; 1541 BP.	
<hr/>			
XX	AC	AAA27924;	
XX	AC		
XX	DT	12-SEP-2000 (first entry)	
XX	DE	Maize Rad2/FEN-1 cDNA.	
XX	DE		
KW	Maize: Rad2/FEN-1, transgenic plant; male sterile plant; endonuclease; exonuclease; DNA repair; gene targeting; ss.		

## RESULT 2

ID AAA27924 standard; cDNA; 1541 BP.

AC AAA27924

DT 12-SEP-2000 (first entry)

Maize Rad2/FEN-1 cDNA

AA Maize; Rad2/FEN-1; transgenic plant; male sterile plant;  
KW endonuclease; exonuclease; DNA repair; gene targeting; ss  
KW





OY 1401 CGCGTTTTTACGTCGCTTGAAGA 1431  
|  
Db 1395 CGCTGTTTTTACTCTTGTCTTGAAGTA 1425

RESULT 3  
AAA27925

ID AAA27925 standard; cDNA; 1381 BP.

AC AAA27925;

DT 12-SEP-2000 (first entry)

DE Maize Rad2/FEN-1 cDNA.

KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;  
endonuclease; exonuclease; DNA repair; gene targeting; ss.

OS Zea mays.

FH Key Location/Qualifiers  
FT CDS 37..1176  
/\*tag= a

PN MO200036109-A1.

PD 22-JUN-2000.

PE 16-NOV-1999; 99WO-US27147.

PR 15-DEC-1998; 98US-0112332.

PA (PION-) PIONEER HI-BRED INT INC.

PI Mahajan PB;

DR WPI: 2000-452026/39.  
P-PSDB; AAY95309.

PT Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA  
recombination and repair in transgenic plants, e.g. for gene targeting  
PT and the production of male sterile plants -

PS Example 1; Page 75-76; 85pp; English.

CC The present sequence is that of maize cDNA coding for RAD2/FEN-1  
(see AAY95309). The cDNA was isolated from a cDNA library prepared  
from W23 line tassel (16-18 cm long) polyA RNA. Rad2/FEN-1 is a  
CC structure specific endonuclease which under certain conditions also  
acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to  
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.  
CC The protein is involved in the regulation of DNA repair and  
recombination in plant systems and therefore may be used for  
CC improving gene targeting during further recombinant DNA protocols  
involving plants. RAD2/FEN-1 endonucleolytic activity is essential  
CC in DNA replication and nucleotide excision and repair reactions.  
CC The exolytic activity is involved in double strand break repair and  
end joining. The protein is also useful in strand exchange  
CC reactions during homologous recombination. These functions may be  
useful in gene targeting and in the production of male sterile  
CC plants. The efficacy of gene targeting can be improved by the  
overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
be produced by the down regulation of Rad2/FEN-1 expression.

SO Sequence 1381 BP; 441 A; 269 C; 346 G; 325 T; 0 other;

Query Match 86.6%; Score 1267; DB 21; Length 1381;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 GCCACAGCCGCCGACAGAGATGGCATCAAGGTTTGACGAACATGCTGGCGGACAA 123  
|  
Db 16 GCCACAGCCGCCGACAGAGATGGCATCAAGGTTTGACGAACATGCTGGCGGACAA 75

OY 124 GGGCCCAAGCGGATGAGAGAGAGAGAGTTCGAGAGTACTTCGGCCGCAAAATGCGCGTC 183  
|  
Db 76 GGGCCCAAGCGGATGAGAGAGAGAGAGTTCGAGAGTACTTCGGCCGCAAAATGCGCGTC 135  
OY 184 GAGCGCAGCATGAGCATATACCACTTCTGATGTAGTTGGAAGAGACAGCGCAAACT 243  
|  
Db 136 GAGCGCAGCATGAGCATATACCACTTCTGATGTAGTTGGAAGAGACAGCGCAAACT 195  
OY 244 CTCACAAATGAAAGCTGTGAAAGTCACTGATGATTTGCAAGGAATGTCACCGGACATA 303  
|  
Db 196 CTCACAAATGAAAGCTGTGAAAGTCACTGATGATTTGCAAGGAATGTCACCGGACATA 255  
OY 304 AGATTACTGSAAGCGGCAATCAAGCCAGTTATATGTTTTGATGCAAGCCCTCTGATATG 363  
|  
Db 256 AGATTACTGSAAGCGGCAATCAAGCCAGTTATATGTTTTGATGCAAGCCCTCTGATATG 315  
OY 364 AAGAAACAAGAGCTTGTCTAAAGATCTCAAAAAGAGATGTCACCAAAAGATCTGACT 423  
|  
Db 316 AAGAAACAAGAGCTTGTCTAAAGATCTCAAAAAGAGATGTCACCAAAAGATCTGACT 375  
OY 424 GAGCGAGTAGAGGTGAGATTAAGATGCGATGAAATGAGCAAGAGGACTGTAAAG 483  
|  
Db 376 GAGCGAGTAGAGGTGAGATTAAGATGCGATGAAATGAGCAAGAGGACTGTAAAG 435  
OY 484 GTACAAAGGCAACACAGAGATTTGTAAGCGCTATTAAAGCTTATGCGGTTCTGT 543  
|  
Db 436 GTACAAAGGCAACACAGAGATTTGTAAGCGCTATTAAAGCTTATGCGGTTCTGT 495  
OY 544 GTAGAGGACCTTCTGAAAGCAGAGACAGATGTCAGCCCTTTGCAATTAACGATAAGTG 603  
|  
Db 496 GTAGAGGACCTTCTGAAAGCAGAGACAGATGTCAGCCCTTTGCAATTAACGATAAGTG 555  
OY 604 TTGCGCTGTGCTTCAGAGATATGAGTCCCTTACTTTGGGGGCCCAAGGTTCTTCTGT 663  
|  
Db 556 TTGCGCTGTGCTTCAGAGATATGAGTCCCTTACTTTGGGGGCCCAAGGTTCTTCTGT 615  
OY 664 CATTTAATGATCAAGTTCCTCAAGAAATATACCTGTGATGATTTGATTTGCCAAGTT 723  
|  
Db 616 CATTTAATGATCAAGTTCCTCAAGAAATATACCTGTGATGATTTGATTTGCCAAGTT 675  
OY 724 TTGAGAGAGCTTGAACCTACCATGACCAAGTTGATTTGTGATCTGTGTGAGAT 783  
|  
Db 676 TTGAGAGAGCTTGAACCTACCATGACCAAGTTGATTTGTGATCTGTGTGAGAT 735  
OY 784 GACTATTGATGATGATCAAGATCAAGGTATCGGGGGCAACAGCTGCAAACTATTGCTCA 843  
|  
Db 736 GACTATTGATGATGATCAAGATCAAGGTATCGGGGGCAACAGCTGCAAACTATTGCTCA 795  
OY 844 CATGGGTCATAGAAAGCAATCTTTGAGAAATCTTAATAAAGACAGATATCAAAATCTTGAG 903  
|  
Db 796 CATGGGTCATAGAAAGCAATCTTTGAGAAATCTTAATAAAGACAGATATCAAAATCTTGAG 855  
OY 904 GACTGGCTTACCAAGAAAGCTGACGCTTTGCAAGAGCCCTAATGACATTTGGATTT 963  
|  
Db 856 GACTGGCTTACCAAGAAAGCTGACGCTTTGCAAGAGCCCTAATGACATTTGGATTT 915  
OY 964 CCGAGCTTAATGAGACCTGACCTGATGAGAGGGTCTCATTAAGTTCTCTGTAAGAT 1023  
|  
Db 916 CCGAGCTTAATGAGACCTGACCTGATGAGAGGGTCTCATTAAGTTCTCTGTAAGAT 975  
OY 1024 AATGTTTTCAAGAAAGATGGGTGACAAAGGCAATGAGAAGATCAATCTTGCAAGAT 1083  
|  
Db 976 AATGTTTTCAAGAAAGATGGGTGACAAAGGCAATGAGAAGATCAATCTTGCAAGAT 1035  
OY 1084 AATGTCGCCAAGGAAGACTCGATGCTTTTTCAGACCCCACTGCGCACCAATGCGCACCG 1143  
|  
Db 1036 AATGTCGCCAAGGAAGACTCGATGCTTTTTCAGACCCCACTGCGCACCAATGCGCACCG 1095  
OY 1144 CTAAAAAGGAGAGAGACTCGATGATTAACCAAGCAAGGACGCTGCGAACAAAGAAACAAG 1203  
|  
Db 1096 CTAAAAAGGAGAGAGACTCGATGATTAACCAAGCAAGGACGCTGCGAACAAAGAAACAAG 1155

OY	1204	GCCTGGTGGAAAGAGAAATAAATCTTGAGCTTGAGTGAATGTAACAACACTGCAGCTACGGAACGAG	1263
OY	1205	CGCTGGTGGAAAGAGAAATAAATCTTGAGCTTGAGTGAATGTAACAACACTGCAGCTACGGAACGAG	1264
Db	1156	GCCTGGTGGAAAGAGAAATAAATCTTGAGCTTGAGTGAATGTAACAACACTGCAGCTACGGAACGAG	1215
OY	1264	CGGTGGCGTGATCAGTCCGCTAGATTATTTAACTCCCTGTGTTTTAACTCAGAGCTTTGGT	1323
OY	1324	AAAAGTT 1330	
Db	1216	CGGTGGCGTGATCAGTCCGCTAGATTATTTAACTCCCTGTGTTTTAACTCAGAGCTTTGGT	1275
OY	1276	AAAAGTT 1282	
XX	RESULT 4		
XX	AAA27926		
XX	AAA27926	standard; cDNA; 1478 BP.	
XX	AAA27926;		
XX	12-SEP-2000	(first entry)	
XX	Maize Rad2/FEN-1	CDNA.	
XX	Maize; Rad2/FEN-1;	transgenic plant; male sterile plant;	
KW	endonuclease; exonuclease; DNA repair;	gene targeting; ss.	
XX	Zea mays.		
XX	Key	Location/Qualifiers	
FT	CDS	97..1236	
FT		/*tag= a	
XX	WO200036109-A1.		
XX	22-JUN-2000.		
XX	16-NOV-1999;	99MO-US27147.	
XX	15-DEC-1998;	98US-0112332.	
XX	(P10N-) PIONEER HI-BRED INT INC.		
XX	Mahaajan PB;		
XX	MP1: 2000-452026/39.		
XX	P-PSDB: AAY95310.		
XX	Example 1; Page 77-79; 85pp; English.		
XX	The present sequence is that of maize cDNA coding for RAD2/FEN-1		
XX	(see AAY95310). The cDNA was isolated from a library prepared		
XX	from B73 line seed vitreous endosperm RNA. Rad2/FEN-1 is a		
XX	structure specific endonuclease which under certain conditions also		
XX	acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to		
XX	produce the Rad2/FEN-1 polypeptides in transgenic plant cells.		
XX	The protein is involved in the regulation of DNA repair and		
XX	recombination in plant systems and therefore may be used for		
XX	improving gene targeting during further recombinant DNA protocols		
XX	involving plants. Rad2/FEN-1 endonucleolytic activity is essential		
XX	in DNA replication and nucleotide excision and repair reactions.		
XX	The exolytic activity is involved in double strand break repair and		
XX	end joining. The protein is also useful in strand exchange		
XX	reactions during homologous recombination. These functions may be		
XX	useful in gene targeting and in the production of male sterile		
XX	plants. The efficacy of gene targeting can be improved by the		
XX	overexpression of exogenous Rad2/FEN-1 while male sterile plants can		
XX	be produced by the down regulation of Rad2/FEN-1 expression.		
XX	Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other;		

Query Match	61.1%;	Score 894;	DB 21;	Length 1478;
Best Local Similarity	99.48;	Pred. No. 0;		
Matches 1244;	Conservative	0;	Mismatches 7;	Indels 0;
				Gaps 0.

QY	20	CGGGTTTTCTTGGCCGACCTCCGGCTCAAGCGCGCGCCGACCGCCGACCGCCGCGAG	79
Db	32	CGCGGGTTTTCTTGGCCGACCTCCGGCTCAAGCGCGCGCCGACCGCCGACCGCCGCGAG	91
QY	80	ACGAGATGGGCATCAAGGGTTTACGCAAACTGCTGGCGGACAAATGCGCCCAAGCGATGA	139
Db	92	ACGAGATGGGCATCAAGGGTTTACGCAAACTGCTGGCGGACAAATGCGCCCAAGCGATGA	151
QY	140	AGGAGCAGAAGTTTCGAGAGCTACTTGGGCGCGCAAAATGCGCGTCCAGCGCAGCATGAGCA	199
Db	152	AGGAGCAGAAGTTTCGAGAGCTACTTGGGCGCGCAAAATGCGCGTCCAGCGCAGCATGAGCA	211
QY	200	TATACCAAGTTCCTGATTTGTAGTTGGGAAGGACAGCGCTGGAAACTCTCCAAATGAAGCTG	259
Db	212	TCTACCAAGTTCCTGATTTGTAGTTGGGAAGGACAGCGCTGGAAACTCTCCAAATGAAGCTG	271
QY	260	GTGAAGTCATCAATGCTTTGTCAGAGAAATGTTCAACCGGCACTTAAGATTACTGCAAGCGG	319
Db	272	GTGAAGTCATCAATGCTTTGTCAGAGAAATGTTCAACCGGCACTTAAGATTACTGCAAGCGG	331
QY	320	GAATCAAGCCAGTTTATGTTTTGATGGGCAAGCGCTCTATATGTAAGAACAAGAGCTTG	379
Db	332	GAATCAAGCCAGTTTATGTTTTGATGGGCAAGCGCTCTATATGTAAGAACAAGAGCTTG	391
QY	380	CTAAAGATTAATCAAAAAGATGATGCAACCAAAATCTGACTGAGGCAGTAAGAGTGA	439
Db	392	CTAAAGATTAATCAAAAAGATGATGCAACCAAAATCTGACTGAGGCAGTAAGAGTGA	451
QY	440	GAGATAAAGATGCATTTGAAAAAAATTGAGCAGAAGAGCTCTAAAGGTCACAAGGCAACACA	499
Db	452	GAGATAAAGATGCATTTGAAAAAAATTGAGCAGAAGAGAGCTCTAAAGGTCACAAGGCAACACA	511
QY	500	ACGAAATTTGTAACCGGCTTTTAAGACTTAATGGGGGTTCTGTGTAAAGAGCCCTTGCG	559
Db	512	ACGAAATTTGTAACCGGCTTTTAAGACTTAATGGGGGTTCTGTGTAAAGAGAGCCCTTGCG	571
QY	560	AAGCAGAAGCGAATGTGCGAGGCCCTTTGGCTAAACGATTAAGGTGTTGCGTGTGCTTACG	619
Db	572	AAGCAGAAGCGAATGTGCGAGGCCCTTTGGCTAAACGATTAAGGTGTTGCGTGTGCTTACG	631
QY	620	AAGATATGGAATCCCTTACTTTTGGGGGCTCACAGGTTCTTGTGATTTAATGATGCCAA	679
Db	632	AAGATATGGAATCCCTTACTTTTGGGGGCTCACAGGTTCTTGTGATTTAATGATGCCAA	691
QY	680	GTTTCCAAAGAAATACCTGTGATGTAATGATTTGCCAAGGTTTTGGAGAGGCTTGTAAC	739
Db	692	GTTTCCAAAGAAATACCTGTGATGTAATGATTTGCCAAGGTTTTGGAGAGGCTTGTAAC	751
QY	740	TCACCATGAGACCAATGTCATTTGATTTTGGCATCTCTGTGTGATGTGACTATTGTGATACGA	799
Db	752	TCACCATGAGACCAATGTCATTTGATTTTGGCATCTCTGTGTGATGTGACTATTGTGATACGA	811
QY	800	TCAAAGGATATGGGGGGCAAAAGAGCTCTCAAACTTATTTGTCAAACATGGGTCATAGAAA	859
Db	812	TCAAAGGATATGGGGGGCAAAAGAGCTCTCAAACTTATTTGTCAAACATGGGTCATAGAAA	871
QY	860	GCATCTTGGACATCTTAATTAAGACAGATATCAAAATTCCTAGAGACTGGCCTTACCAAG	919
Db	872	GCATCTTGGACATCTTAATTAAGACAGATATCAAAATTCCTAGAGACTGGCCTTACCAAG	931
QY	920	AAGTCGACGCTTTGTCAGAGAGCTATGTCACATGATATTTCTGAGCTAAATATGCA	979
Db	932	AAGTCGACGCTTTGTCAGAGAGCTATGTCACATGATATTTCTGAGCTAAATATGCA	991
QY	980	CTGCACTGATGAGAGAGGTCATATAAGTTTCTGGTTAAAGATAATGTTTCAACAGAG	1039
Db	992	CTGCACTGATGAGAGAGGTCATATAAGTTTCTGGTTAAAGATAATGTTTCAATGTAAG	1051

QY 1040 ATCGGGTGAACAAGGCCATGAGAGATCAATCTGCGCAAGAAATTAATGTCGCAAGGAA 1099  
 |||||||  
 Db 1052 ATCGGGTGAACAAGGCCATGAGAGATCAATCTGCGCAAGAAATTAATGTCGCAAGGAA 1111  
 |||||||  
 QY 1100 GACTCGAGTCTTTTTCACAGCCAACTGCGCACCATCATCAGCCGCTAAACGGAAGAGA 1159  
 |||||||  
 Db 1112 GACTCGAGTCTTTTTCACAGCCAACTGCGCACCATCATCAGCCGCTAAACGGAAGAGA 1171  
 |||||||  
 QY 1160 CTTCGGATTAACAAGCAAGCGAGCTGCGCAAGAAAGAAAGCTGCTGGAAGAGAGA 1219  
 |||||||  
 Db 1172 CTTCGGATTAACAAGCAAGCGAGCTGCGCAAGAAAGAAAGCTGCTGGAAGAGAGA 1231  
 |||||||  
 QY 1220 AATAATCTTGATGCTGATGATCAACTAGACTAGCAAGCAAGCGGTCGC 1270  
 |||||||  
 Db 1232 AATAATCTTGATGCTGATGATCAACTAGACTAGCAAGCAAGCGGTCGC 1282  
 |||||||

## RESULT 5

AAFI5908

ID AAFI5908 standard; cDNA; 584 BP.

XX AAFI5908;

XX 13-MAR-2001 (first entry)

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:343.

XX Human: prostate cancer; prostate cancer antigen; detection; diagnosis;  
 XX neuroprotective; cytosolic; cardiovascular; immunomodulatory; muscular;  
 XX vulnary; gastrointestinal; nephrologic; antiinfective; gynaecological;  
 XX antibacterial; gene therapy; neural; immune; reproductive; renal;  
 XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
 XX wound; infectious disease; ss.

OS Homo sapiens.

XX WO200055174-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000MO-US05988.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587513/55.

XX P-PSDB; AAB56705.

PT Prostate cancer associated gene sequences, referred to as prostate  
 PT cancer antigens, useful for treatment, prevention, and diagnosis of  
 PT disorders such as prostate cancer -

XX Claim 1; Page 846; 2338bp; English.

XX AAFI5566 to AAFI6505 encode the human prostate cancer associated  
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
 CC The prostate cancer antigens can have neuroprotective, cytosolic,  
 CC cardiovascular, immunomodulatory, muscular, vulnary, gastrointestinal,  
 CC nephrologic, antiinfective, gynaecological and antibacterial activities,  
 CC and can be used in gene therapy. The prostate cancer antigen  
 CC polynucleotides may be used for detection of prostate cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The prostate cancer antigens may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to  
 CC AAB57303 represent sequences used in the exemplification of the present  
 CC invention.

SQ Sequence 584 BP; 188 A; 120 C; 110 G; 165 T; 1 other;

Query Match 2.7%; Score 40; DB 21; Length 584;

Best Local Similarity 100.0%; Pred. No. 4.7e-06;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1424 TTGAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1463  
 |||||||Db 543 TTGAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 582  
 |||||||

## RESULT 6

AAH69821/C

ID AAH69821 standard; cDNA; 288 BP.

XX AAH69821;

XX 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 1095.

XX Cervical cancer; cytosolic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

XX WO200142467-A2.

XX 14-JUN-2001.

XX 08-DEC-2000; 2000MO-US3312.

XX 08-DEC-1999; 99US-0166681.

XX 21-DEC-1999; 99US-0171350.

XX 14-MAR-2000; 2000US-0189315.

XX 12-MAY-2000; 2000US-0203791.

XX 09-JUN-2000; 2000US-0210600.

XX 21-JUL-2000; 2000US-0220114.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Deeds J, Berger A, Zhao X;

XX WPI; 2001-375006/39.

PT New isolated nucleic acid for diagnosing and treating cervical cancer  
 PT and for assessing and detecting compounds for treating the cancer -

XX Claim 1; Page 279; 1051bp; English.

CC The invention relates to novel genes (AAH68727-AAH73383) associated with  
 CC cervical cancer with cytosolic activity. The nucleic acids and encoded  
 CC polypeptides are useful: to assess if a patient is afflicted with  
 CC cervical cancer or has a pre-malignant condition; to monitor the  
 CC progression of cervical cancer or a premalignant condition in a patient;  
 CC and to select and/or assess the efficacy of a compound or therapy for  
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
 CC useful for gene therapy.

XX Sequence 288 BP; 88 A; 49 C; 42 G; 107 T; 2 other;

XX Query Match 2.7%; Score 39; DB 22; Length 288;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1423 TTGGAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1461  
 |||||||Db 66 TTGGAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 28  
 |||||||

## RESULT 7

AAH71246/C

ID AAH71246 standard; cDNA; 305 BP.

XX AAH71246 standard; cDNA; 305 BP.



PS Claim 1: Page 593; 1425bp; English.  
XX  
CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
CC associated proteins and polynucleotide sequences, their agonists, and  
CC antagonists may have neuroprotective; cytostatic; cardioactive; and  
CC immunomodulatory; muscular active general; vulnery; gastrointestinal  
CC general; nephrotropic; anti-infective; gynecological; or antibacterial  
CC activity. The invention also includes antibodies specific for the  
CC protein or polynucleotide sequences. The lung cancer associated  
CC polynucleotide sequences may be used for detection of lung cancer,  
CC chromosome identification, as chromosome markers, and for numerous other  
CC diagnostic or research purposes. The proteins may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders. The proteins may also be used in the treatment of wounds and  
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
CC peptide AAB58549 are used in the course of the invention for the  
CC identification and characterisation of the polynucleotide and protein  
CC sequences.  
XX  
SQ Sequence 2381 BP; 527 A; 704 C; 659 G; 486 T; 5 other;  
XX  
Query Match 2.6%; Score 38; DB 21; Length 2381;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 2272 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2309  
RESULT 10  
AAS26809  
ID AAS26809 standard; DNA; 8711 BP.  
XX  
AC AAS26809;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
XX Human genomic DNA encoding partial novel secreted protein, Seq ID 1783.  
DE  
XX  
XX Human; immunosuppressive; antiarthritic; ds; antipneumatic;  
KW cytosolic; cardiant; vasotropic; cerebroprotective; nocrotic;  
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
KW vulnery; secreted protein; rheumatoid arthritis;  
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
KW cerebrovascular disorder; cerebral ischemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
KW corneal infection; wound healing; epithelial cell proliferation;  
KW skin ageing; food additive; preservative; antiproliferative.  
XX  
OS Homo sapiens.  
XX  
XX  
XX WO200155322-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01341.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0186874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225457.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226682.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234274.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0235484.  
PR 26-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 27-SEP-2000; 2000US-0235837.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241765.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.



CC treatment and diagnosis of colon disorders, such as colon cancer. The  
CC polynucleotides may be used in diagnostics and research, such as for  
CC chromosome identification, and as hybridisation probes. The proteins  
CC may also be used to prevent diseases such as neural disorders, immune  
CC system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, wounds, renal disorders, infectious  
CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and  
CC AAB54007 represent sequences used in the exemplification of the present  
CC invention.  
CC  
CC  
CC Sequence 302 BP; 159 A; 20 C; 35 G; 86 T; 2 other;  
XX  
SQ  
Query Match 2.5%; Score 37; DB 21; Length 302;  
Best Local Similarity 100.0%; Pred. NO. 7.4e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 207 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 243  
RESULT 12  
AAI83214  
ID AAI83214 standard; cDNA: 342 BP.  
XX  
AC AAI83214;  
XX  
XX 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 3274.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KM nervous system disorders; arthritis; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
XX 26-FEB-2001; 2001WO-US04927.  
XX  
PF 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-514838/56.  
DR P-PSDB; AAO03283.  
XX  
XX Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
XX  
PS Claim 1; SEQ ID NO 3274; 1399pp + Sequence Listing; English.  
XX  
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulatory  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 342 BP; 149 A; 52 C; 74 G; 67 T; 0 other;  
SQ  
Query Match 2.5%; Score 37; DB 22; Length 342;  
Best Local Similarity 100.0%; Pred. NO. 7.2e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 54 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 90  
RESULT 13  
AAF94934/C  
ID AAF94934 standard; cDNA: 396 BP.  
XX  
XX AAF94934;  
XX  
XX 23-MAY-2001 (first entry)  
XX  
DE Human ovarian cancer associated coding sequence SEQ ID NO: 125.  
XX  
XX Human; ovarian cancer; vaccine; gene therapy; carcinoma; ss.  
KW  
XX  
OS Homo sapiens.  
XX  
PN WO200118046-A2.  
XX  
PD 15-MAR-2001.  
XX  
XX 08-SEP-2000; 2000WO-US24827.  
XX  
PF 10-SEP-1999; 99US-0394374.  
PR 01-MAY-2000; 2000US-0561178.  
PR 15-AUG-2000; 2000US-0640173.  
PR 07-SEP-2000; 2000US-0656668.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Xu J, Stolk JA;  
XX  
XX WPI; 2001-211395/21.  
XX  
XX Isolated polypeptides associated with ovarian carcinomas, and the  
PT nucleic acids that encode them, useful for the prevention diagnosis and  
PT treatment of ovarian cancers -  
XX  
XX  
PS Claim 5; Page 157; 189pp; English.  
XX  
XX The present invention provides a number of coding sequences and proteins,  
CC the over-expression of which is associated with ovarian carcinoma/cancer.  
CC These can be used in the diagnosis, treatment and prevention of ovarian  
CC cancer, optionally by gene therapy or in the form of a vaccine. The  
CC present sequence is an example of one of these sequences.  
XX  
XX  
XX Sequence 396 BP; 115 A; 71 C; 60 G; 125 T; 25 other;  
XX  
XX  
XX Query Match 2.5%; Score 37; DB 22; Length 396;  
Best Local Similarity 100.0%; Pred. NO. 7.1e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 40 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4  
RESULT 14  
AAI88595  
ID AAI88595 standard; cDNA: 411 BP.  
XX  
XX AAI88595;  
XX



DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 8655.  
XX  
XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
XX  
PR 18-MAY-2000; 2000US-0577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI: 2001-514838/56.  
DR P-PSDB: AAO08664.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
PS Claim 1; SEQ ID NO 8655; 1399pp + Sequence Listing; English.  
XX  
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 411 BP; 214 A; 24 C; 64 G; 107 T; 2 other;  
XX  
Query Match 2.5%; Score 37; DB 22; Length 411;  
Best Local Similarity 100.0%; Pred. No. 7e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 156 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 192  
XX  
RESULT 15  
AAI83922  
ID AAI83922 standard; cDNA; 412 BP.  
XX  
AC AAI83922;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 3982.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.

XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
XX  
PR 18-MAY-2000; 2000US-0577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI: 2001-514838/56.  
DR P-PSDB: AAO03991.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
PS Claim 1; SEQ ID NO 3982; 1399pp + Sequence Listing; English.  
XX  
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 412 BP; 161 A; 44 C; 96 G; 104 T; 7 other;  
XX  
Query Match 2.5%; Score 37; DB 22; Length 412;  
Best Local Similarity 100.0%; Pred. No. 7e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 113 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 149  
XX  
RESULT 16  
AAL19865/c  
ID AAL19865 standard; cDNA; 426 BP.  
XX  
AC AAL19865;  
XX  
DT 07-DEC-2001 (first entry)  
XX  
DE Human breast cancer expressed polynucleotide 12322.  
XX  
KW Human; breast cancer; cell marker; cytosaltic; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200151628-A2.  
XX  
PD 19-JUL-2001.  
XX  
PF 10-JAN-2001; 2001WO-US00798.  
XX  
PR 14-JAN-2000; 2000US-0176077.  
PR 14-MAR-2000; 2000US-0189167.

PR 24-MAR-2000; 2000US-0192099.  
PR 29-MAR-2000; 2000US-0193480.  
PR 15-MAY-2000; 2000US-0205230.  
PR 09-JUN-2000; 2000US-0211315.  
PR 25-JUL-2000; 2000US-0220534.  
XX  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PI Lillie J, Xu Y, Wang Y, Steilmann K;  
XX WPI; 2001-451856/48.  
DR  
XX  
XX  
XX New peptide useful as a marker for the diagnosis of breast cancer -  
PS  
XX  
XX Claim 1; Page 2177-2178; 3695pp; English.  
XX  
XX The invention relates to human breast cancer expressed polynucleotides  
CC (AAID07544-AAI26789) and methods of assessing whether a patient is  
CC afflicted with breast cancer by examining the correlation between the  
CC expression of certain markers and the cancerous state of breast cells.  
CC The polynucleotides and encoded polypeptides are potential markers for  
CC detecting, diagnosing, monitoring, characterising treating and  
CC potentially preventing breast cancer. The polynucleotides and encoded  
CC polypeptides are also useful for isolating compounds with cytostatic  
CC activity.  
XX  
XX Sequence 426 BP; 125 A; 56 C; 74 G; 171 T; 0 other;  
SQ  
Query Match 2.5%; Score 37; DB 22; Length 426;  
Best Local Similarity 100.0%; Pred. No. 7e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 47 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 11  
RESULT 17  
AAA96571  
ID AAA96571 standard; DNA; 524 BP.  
AC  
XX AAA96571;  
AC  
XX  
XX 08-FEB-2001 (first entry)  
DE Noncoding region of core 2 beta-1,6-N-acetylglucosaminyltransferase gene.  
XX  
XX  
XX Core 2 beta-1,6-N-acetylglucosaminyltransferase; core2b GlcNAc-T;  
KW cancer; cardiovascular disorder; inflammatory disorder; asthma;  
KW rheumatoid arthritis; inflammatory bowel disease; arteriosclerosis;  
KW septic shock; adult respiratory distress syndrome; ARDS; cancer;  
KW platelet-mediated pathology; atherosclerosis; gastrointestinal disorder;  
KW clotting; ascites; cholelithiasis; cirrhosis; Crohn's disease;  
KW diverticulitis; ulcerative colitis; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX CA2296936-A1.  
PN  
XX  
XX 03-AUG-2000.  
PD  
XX  
XX 03-FEB-2000; 2000CA-2296936.  
PF  
XX  
XX 03-FEB-1999; 99US-0118674.  
PR  
XX  
XX (GLYC-) GLYCDESIGN INC.  
PA  
XX  
XX Korczak B, Lew A;  
PI  
XX  
XX WPI; 2000-594746/57.  
DR  
XX  
XX New nucleic acid molecules of core 2  
PT beta-1,6-N-acetylglucosaminyltransferase useful for providing new

PT compositions for treatment of disorders mediated by the enzyme  
PT including cancer, cardiovascular and inflammatory disorders. -  
XX  
XX  
XX Disclosure; Page 52; 66pp; English.  
PS  
XX  
XX The present sequence represents a noncoding region from a human core 2  
CC beta-1,6-N-acetylglucosaminyltransferase (core2b GlcNAc-T) gene.  
CC The polypeptide can be used to treat diseases and disorders, such as  
CC cancer, cardiovascular disorders and inflammatory disorders including  
CC asthma, rheumatoid arthritis, inflammatory bowel disease,  
CC arteriosclerosis, septic shock, adult respiratory distress syndrome  
CC (ARDS) and cancer. Various platelet-mediated pathologies such as  
CC atherosclerosis and clotting can also be treated. The polypeptides of  
CC the invention are predominantly expressed in gastrointestinal tissue  
CC (stomach, colon, intestine, testis) and are elevated in cancer.  
CC Gastrointestinal disorders that may be prevented or treated include  
CC ascites, cholelithiasis, cirrhosis, Crohn's disease, diverticulitis  
CC and ulcerative colitis. The antibodies may be used in  
CC immuno-histochemical analysis, to detect the novel polypeptide and to  
CC localize it to particular cells and tissues and to specific subcellular  
CC locations and to quantitate the level of expression.  
XX  
XX Sequence 524 BP; 171 A; 90 C; 127 G; 136 T; 0 other;  
SQ  
Query Match 2.5%; Score 37; DB 21; Length 524;  
Best Local Similarity 100.0%; Pred. No. 6.7e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 452 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 488  
RESULT 18  
AAf64213  
ID AAF64213 standard; cDNA; 699 BP.  
AC  
XX AAF64213;  
AC  
XX  
XX 06-APR-2001 (first entry)  
DE Human secreted protein gene 38 SEQ ID NO:48.  
XX  
XX  
XX Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KW neurotropic; neuroprotective; antibacterial; vituicide; fungicide;  
KW ophthalmological; vulnerrary; autoimmune disease; cardiovascular disorder;  
KW hyperproliferative disorders; cerebrovascular disorder; wound healing;  
KW nervous system disorder; ocular disorder; skin aging; chemotaxis;  
XX  
XX food additive; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200077026-A1.  
PN  
XX  
XX 21-DEC-2000.  
PD  
XX  
XX 01-JUN-2000; 2000WO-US14973.  
PF  
XX  
XX 11-JUN-1999; 99US-0138630.  
PR  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX (ROSE/) ROSEN C A.  
PI  
XX  
XX Rosen CA, Ruben SM, Komatsoulis GA;  
PI  
XX  
XX WPI; 2001-071258/08.  
DR  
XX  
XX P-PSDB; AAB75543.  
DR  
XX  
XX Nucleic acid molecules encoding human secreted proteins, used in  
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
PT Parkinson's diseases and cancers -  
XX



```

XX Claim 4: Pages 5-6; 10pp; Japanese.
XX
XX AAQ79736 encodes AAR66604 a flower style-specific S-ribonuclease,
CC isolated from a self-incompatible wild tomato species.
XX
XX Sequence 826 BP; 309 A; 150 C; 131 G; 236 T; 0 other:
SQ
Query Match 2.5%; Score 37; DB 16; Length 826;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1427 AACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 774 AACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810

RESULT 21
AAT64553
ID AAT64553 standard; cDNA to mRNA; 826 BP.
XX
AC AAT64553;
XX
XX 22-MAY-1997 (first entry)
DT
XX Tomato S-ribonuclease cDNA fragment.
DE
XX S-ribonuclease; plant style; promoter; tissue-specific;
KW self-incompatibility; ds.
XX
XX Lycopersicon peruvianum.
OS
XX
XX Key Location/Qualifiers
FH CDS 1..489
FT /tag= a
XX
XX JP09028381-A.
XX
XX 04-FEB-1997.
PD
XX
XX 24-JUL-1995; 95JP-0187557.
PE
XX
XX 24-JUL-1995; 95JP-0187557.
PR
XX
XX (CHEN/) CHENG E. K.
XX (MINA/) MINAMI H.
XX (TAKA/) TAKAGI M.
XX
XX WPI: 1997-159091/15.
DR P-PSDB; AAM14912.
XX
XX
XX Tomato S-ribonuclease gene and promoter sequence - useful for
XX expressing genes in plant style cells, e.g. to confer
XX self-incompatibility on the plant
XX
XX Claim 4; Page 10-11; 15pp; Japanese.
XX
XX The present sequence encodes part of a tomato (Lycopersicon
XX peruvianum) S-ribonuclease. The S-gene can confer self-
XX incompatibility on a plant.
XX
XX Sequence 826 BP; 309 A; 150 C; 131 G; 236 T; 0 other:
SQ
Query Match 2.5%; Score 37; DB 18; Length 826;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1427 AACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 774 AACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810

RESULT 22

```

```

AAQ04441
ID AAQ04441 standard; DNA; 957 BP.
XX
XX AAQ04441;
AC
XX 26-SEP-1989 (first entry)
DT
XX
XX Aequorin gene.
DE
XX
XX Aequorin protein; aequorin; calcium levels; ss.
KW
XX
XX Aequoria victoria.
OS
XX
XX Key Location/Qualifiers
FH CDS 115..702
FT /tag= a
FT /label=aequorin
XX
XX JP02096597-A.
XX
XX 09-APR-1990.
PD
XX
XX 24-AUG-1984; 84JP-0113000.
PE
XX
XX 01-JAN-1989; 89JP-0176125.
PR
XX
XX (CHCC ) CHITSO CORP.
XX
XX WPI: 1990-151860/20.
XX P-PSDB; AAR04822.
XX
XX Biosynthesis of photoprotein aequorin - comprises culturing modified host
XX contg. specific DNA sequence of photoprotein aequorin
XX
XX Disclosure; ; Japanese.
XX
XX A prokaryotic host is transformed with this sequence in the prodn. of the
XX photoprotein aequorin. The protein is useful for analysing low calcium
XX levels.
XX
XX Sequence 957 BP; 362 A; 156 C; 211 G; 228 T; 0 other:
SQ
Query Match 2.5%; Score 37; DB 11; Length 957;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 890 AACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 926

RESULT 23
AAC92233
ID AAC92233 standard; cDNA; 958 BP.
XX
XX AAC92233;
AC
XX
XX 21-MAR-2001 (first entry)
DT
XX
XX Apoaequorin-encoding gene nucleotide sequence SEQ ID NO.5.
DE
XX
XX Bioluminescence generating system; fluorescent protein; recreation;
XX entertainment; amusement; toy; paint; slimy play material; textile;
XX bubble making toy; balloon; cosmetic; bath powder; body lotion; gel;
XX powder; cream; toothpaste; dentifrice; soap; body paint; bubble bath;
XX gelatin; icing; frosting; beverage; beer; wine; champagne; soft drink;
XX glowing ice; fountain; liquid firework; ss.
XX
XX Unidentified.
XX
XX US6152358-A.
XX
XX 28-NOV-2000.
PD

```

[illegible]

FT	XX	/product= "Aequorea victoria aequorin protein"
XX	WO200168824-A2.	
XX	20-SEP-2001.	
XX	15-MAR-2001; 2001WO-US08277.	
XX	15-MAR-2000; 2000US-189691P.	
XX	(PROL-) PROLUME LTD.	
XX	(BRIA/) BRIAN B J.	
XX	Bryan BJ, Szent-Gyorgyi C, Szczepaniak W;	
XX	MP1: 2002-010561/01.	
XX	P-PSDB; AAEL3370.	
XX	Nucleic acids encoding Renilla reniformis green fluorescent proteins,	
XX	useful in diagnostic bioluminescence procedures -	
XX	Disclosure; Page 146-147; 175pp; English.	
XX	The patent discloses sea pansy (Renilla reniformis) green fluorescent	
XX	proteins (GFP) and their corresponding polynucleotides. The invention	
XX	also relates to sequences of the bioluminescence generating system	
XX	(e.g. luciferase). R. reniformis GFP are used in diagnostic methods	
XX	and in the production of novelty items such as toys (e.g. squirt gun,	
XX	pellet gun, toy "Halloween" eggs, toy cigarettes, board/card game	
XX	toys), finger paints, slimy play material, bubbles in bubble making	
XX	toys, fishing lures, dolls, sparklers, magic wand toys, balloons,	
XX	personal care item (e.g. cosmetic, bath powders, body creams, tooth	
XX	pastes, mouth wash, soaps, body paints, bubble bath), ornamental	
XX	transgenic plants, fountain, fairy dust, food (gelatins, icings,	
XX	frosting, beers, wine, champagne, milk, soft drinks, ice cubes, ice,	
XX	dry ice, beverage), textile (foot bag, clothing) and/or paper product	
XX	(greeting cards, wrapping paper). The present sequence is a DNA	
XX	encoding Aequorea victoria aequorin protein.	
XX	Sequence 958 BP; 363 A; 156 C; 210 G; 229 T; 0 other;	
XX	Query Match 2.5%; Score 37; DB 24; Length 958;	
XX	Best Local Similarity 100.0%; Pred. No. 6e-05;	
XX	Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
XX	0Y 1427 AAGTAAAAA	
XX	DB 891 AAGTAAAAA	
XX	RESULT 25	
XX	AAAN81534	
XX	ID AAAN81534 standard; DNA; 959 BP.	
XX	AAAN81534:	
XX	AC 05-DEC-1990 (first entry)	
XX	DE PAQ440 aequorin gene variant.	
XX	KM PAQ440 aequorin; variant; photogenic mechanism; E.coli.	
XX	OS Aequorea.	
XX	Key Location/Qualifiers	
XX	FT old_sequence 220..	
XX	FT /tag= a	
XX	FT /note="G to C"	
XX	FT old_sequence 238..	
XX	FT /tag= b	
XX	FT /note="G to A"	
XX	FT 307..	
XX	FT /tag= c	



XX AC AAS02547;  
XX DT 18-JUL-2001 (first entry)  
XX XX  
XX DE Human secreted protein gene #28.  
XX  
XX KW Human secreted protein; diagnosis; autoimmune disease; ss;  
XX KW Rheumatoid arthritis; hyperproliferative disorder; neoplasm; sunburn;  
XX KW Cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
XX KW Cerebral ischemia; angiogenesis; nervous system disorder; skin aging;  
XX KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
XX KW Wound healing; epithelial cell proliferation; chemotaxis; preservative;  
XX KW organ transplantation; tissue regeneration; food additive.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200123409-A2.  
XX  
XX PD 05-APR-2001.  
XX  
XX PF 26-SEP-2000; 2000WO-US26371.  
XX  
XX PR 27-SEP-1999; 99US-0155804.  
XX  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX PI Rosen CA, Ruben SM, Komatsoulis GA;  
XX  
XX DR WPI: 2001-266139/27.  
XX  
XX DR P-PSDB; AAU01686.  
XX  
XX PT Nucleic acids encoding 38 human secreted polypeptides, useful for  
XX PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's  
XX PT disease and diabetic retinopathy -  
XX  
XX PS Disclosure: Page 436-437; 488pp; English.  
XX  
XX CC AAS02511-AAS02557 represent human secreted protein coding sequences  
XX CC and primers of the invention. The human secreted protein sequences are  
XX CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
XX CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are  
XX CC also used in diagnosing a pathological condition or susceptibility to a  
XX CC pathological condition. The antibodies to human secreted proteins can  
XX CC also be used in alleviating symptoms associated with the disorders and in  
XX CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
XX CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated  
XX CC include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
XX CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
XX CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
XX CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
XX CC infections caused by bacteria, viruses and fungi and ocular disorders  
XX CC e.g. corneal infection. The polypeptides can also be used to aid wound  
XX CC healing and epithelial cell proliferation, to prevent skin aging due to  
XX CC sunburn, to maintain organs before transplantation, for supporting cell  
XX CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
XX CC polypeptides can also be used as a food additive or preservative to  
XX CC increase or decrease storage capabilities.  
XX  
XX SQ Sequence 1159 BP; 395 A; 201 C; 194 G; 369 T; 0 other;

Query Match 2.5%; Score 37; DB 22; Length 1159;  
Best Local Similarity 100.0%; Pred. No. 5.8e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
DB 1032 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1068

RESULT 28  
AAS29087  
ID AAS29087 standard; cDNA; 1252 BP.

XX AC AAS29087;  
XX DT 21-NOV-2001 (first entry)  
XX XX  
XX DE cDNA encoding for human DNA-binding protein #58.  
XX  
XX KW Human: DNA-binding protein; histone; chromo domain protein;  
XX KW Chromatin organisation modifier; Y-box binding protein;  
XX KW DNA organisation; gene transcription; malignant disease;  
XX KW autoimmune disorder; rheumatic disease; genetic abnormality;  
XX KW infectious disease; neurological disorder; gene therapy;  
XX KW immunomodulatory; anti-HIV; anti rheumatic; anti microbial;  
XX KW cytostatic; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO20015162-A1.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US01305.  
XX  
XX PR 31-JAN-2000; 2000US-0179065.  
XX PR 04-FEB-2000; 2000US-0180628.  
XX PR 24-FEB-2000; 2000US-0184664.  
XX PR 02-MAR-2000; 2000US-0186350.  
XX PR 16-MAR-2000; 2000US-0189874.  
XX PR 17-MAR-2000; 2000US-0190076.  
XX PR 18-APR-2000; 2000US-0198123.  
XX PR 19-MAY-2000; 2000US-0205515.  
XX PR 07-JUN-2000; 2000US-0209467.  
XX PR 28-JUN-2000; 2000US-0214886.  
XX PR 30-JUN-2000; 2000US-0215135.  
XX PR 07-JUL-2000; 2000US-0216647.  
XX PR 07-JUL-2000; 2000US-0216880.  
XX PR 11-JUL-2000; 2000US-0217487.  
XX PR 11-JUL-2000; 2000US-0217496.  
XX PR 14-JUL-2000; 2000US-0218290.  
XX PR 26-JUL-2000; 2000US-0220963.  
XX PR 26-JUL-2000; 2000US-0220964.  
XX PR 14-AUG-2000; 2000US-0224518.  
XX PR 14-AUG-2000; 2000US-0224519.  
XX PR 14-AUG-2000; 2000US-0225213.  
XX PR 14-AUG-2000; 2000US-0225214.  
XX PR 14-AUG-2000; 2000US-0225266.  
XX PR 14-AUG-2000; 2000US-0225267.  
XX PR 14-AUG-2000; 2000US-0225268.  
XX PR 14-AUG-2000; 2000US-0225270.  
XX PR 14-AUG-2000; 2000US-0225477.  
XX PR 14-AUG-2000; 2000US-0225757.  
XX PR 14-AUG-2000; 2000US-0225758.  
XX PR 14-AUG-2000; 2000US-0225759.  
XX PR 18-AUG-2000; 2000US-0226279.  
XX PR 22-AUG-2000; 2000US-0226581.  
XX PR 22-AUG-2000; 2000US-0226588.  
XX PR 22-AUG-2000; 2000US-0227182.  
XX PR 23-AUG-2000; 2000US-0227009.  
XX PR 30-AUG-2000; 2000US-0228924.  
XX PR 01-SEP-2000; 2000US-0229287.  
XX PR 01-SEP-2000; 2000US-0229343.  
XX PR 01-SEP-2000; 2000US-0229344.  
XX PR 01-SEP-2000; 2000US-0229345.  
XX PR 05-SEP-2000; 2000US-0229509.  
XX PR 05-SEP-2000; 2000US-0229513.  
XX PR 06-SEP-2000; 2000US-0230437.  
XX PR 06-SEP-2000; 2000US-0230438.  
XX PR 08-SEP-2000; 2000US-0231242.  
XX PR 08-SEP-2000; 2000US-0231243.  
XX PR 08-SEP-2000; 2000US-0231244.  
XX PR 08-SEP-2000; 2000US-0231413.  
XX PR 08-SEP-2000; 2000US-0231414.  
XX PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0232403.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 21-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239335.  
 PR 13-OCT-2000; 2000US-0239337.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 BA  
 FI Rosen CA, Barash SC, Ruben SM;  
 F1  
 XX WPI: 2001-465557/50.  
 DR  
 DR P-PSDB; AAU18211.  
 XX  
 XX Nucleic acid molecules encoding human secreted chromosomal binding  
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.  
 PT Alzheimer's and Parkinson's diseases and cancers -  
 XX  
 XX Claim 4; SEQ ID No 68; 561pp; English.  
 PS  
 XX The present invention relates to the isolation of novel DNA-binding  
 CC proteins (AAU18154-AAU18281), and cDNA and genomic sequences encoding  
 CC for these proteins. DNA-binding proteins such as histones, chromo  
 CC (chromatin organisation modifier) domain proteins, and Y-box binding  
 CC proteins may contribute to diseases resulting from aberrant DNA  
 CC organisation and/or gene transcription. The sequences of the invention  
 CC are useful in screening assays to identify antagonists and/or agonists  
 CC that may enhance or block activities mediated by DNA-binding proteins.  
 CC Blockers of DNA-binding proteins may be useful in treating disorders  
 CC such as malignant diseases (e.g. cancer), autoimmune disorders  
 CC (e.g. diabetes mellitus), rheumatic diseases (e.g. rheumatoid  
 CC arthritis), genetic abnormalities (e.g. cystic fibrosis), infectious  
 CC diseases (e.g. HIV) and neurological disorders (e.g. Alzheimer's  
 CC disease). The polynucleotide sequences of the invention may also be  
 CC used in gene therapy. AAS29030-AAS29157 represent cDNA sequences  
 CC encoding for novel DNA-binding proteins.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from Wipo  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 1252 BP; 410 A; 221 C; 248 G; 373 T; 0 other;  
 Query Match 2.5%; Score 37; DB 22; Length 1252;  
 Best Local Similarity 100.0%; Pred. No. 5; 8e-05;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 DB 1211 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1247  
 RESULT 29  
 ID AAS63048 standard; cDNA; 1293 BP.  
 AC AAS63048;  
 XX  
 XX 29-JAN-2002 (first entry)  
 DT  
 DE Cell death protective sequence CNI-00727.  
 XX  
 XX Human; protective sequence; cell death; cerebral oedema; infection;  
 KW meningitis; degenerative disease; Alzheimer's disease; heart disease;  
 KW motor neuron disease; demyelinating disease; multiple sclerosis; asthma;  
 KW nutritional condition; peripheral nervous system disorder; ischaemia;  
 KW diabetic neuropathy; autoimmune haemolytic anaemia; respiratory system;





Query Match 2.5%; Score 37; DB 19; Length 1362;  
Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 1326 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1362

## RESULT 31

AAFP98437 standard; cDNA; 1362 BP.

AAFP98437;

07-JUN-2001 (first entry)

Human cDNA clone CCI82\_1 sequence SEQ ID 103.

Human; secreted protein; nutrient; cytokine modulator; proliferation;  
differentiation; immune system modulator; tissue growth; chemotactic;  
hemostatic; thrombolytic; anti-inflammatory; tumour inhibition; ss;  
haematopoiesis.

Homo sapiens.

WO200119988-A1.

22-MAR-2001.

14-SEP-2000; 2000WO-US25135.

17-SEP-1999; 99US-039829.

(GEMV) GENETICS INST INC.

Jacobs K, McCoy JM, Lavallee ER, Collins-Racie LA, Evans C;  
Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;

WPI; 2001-244801/25.

P-PSDB; AAB90706.

Isolated nucleic acids encoding polypeptides, useful for modulating  
e.g. cytokine and cell proliferation/differentiation activity, the  
immune system and hematopoiesis regulating activity.

Disclosure; Page 442; 557pp; English.

Human cDNA clones represented in AAF98374 - AAF98489 encode secreted  
proteins AAB90667 - AAB90750. The cDNA clones are isolated from various  
tissue types, and may be used in the prevention, treatment and diagnosis  
of diseases associated with inappropriate protein expression. The  
polypeptides and nucleic acids may be used as nutrients or to modulate  
cytokine and cell proliferation/differentiation activity and may also be  
involved in modulation of the immune system. The cDNA sequences,  
proteins, their agonists and/or antagonists exhibit haematopoiesis  
regulating activity; tissue growth activity; activity/inhibit activity;  
chemotactic/chemokinetic activity; haemostatic and thrombolytic  
activity; receptor/ligand activity; anti-inflammatory activity;  
hematopoiesis activity; cachectin/tumour suppressor activity; and/or  
tumour inhibition activity. Included in the invention are probes  
represented in AAF98490 - AAF98572 which are specific for the cDNA clones  
encoding the secreted proteins.

Sequence 1362 BP; 409 A; 210 C; 197 G; 544 T; 2 other:

Query Match 2.5%; Score 37; DB 22; Length 1362;

Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463

DB 1326 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1362

## RESULT 32

AAFP33111 standard; cDNA; 1380 BP.

AAFP33111;

23-MAR-2001 (first entry)

Human secreted protein gene 17 SEQ ID NO:27.

Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
neurotropic; neuroprotective; antibacterial; vincicide; fungicide;  
ophthalmological; autoimmune disease; hyperproliferative disorder;  
cardiovascular disorder; cerebrovascular disorder; infection; chemotaxis;  
nervous system disorder; ocular disorder; skin aging; wound healing;  
food additive; tissue regeneration; ss.

Homo sapiens.

WO200077256-A1.

21-DEC-2000.

01-JUN-2000; 2000WO-US14963.

11-JUN-1999; 99US-0138631.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Komatsoulis GA;

WPI; 2001-032315/04.

P-PSDB; AAB64789.

Isolated nucleic acid molecule encoding a human secreted protein is  
used in preventing, treating or ameliorating a medical condition -  
Claim 1; Page 437-438; 506pp; English.

polynucleotide sequences AAF33095 - AAF33142 encode human secreted  
proteins AAB64773 - AAB64820. Fragments of the secreted proteins and  
amino acid sequences which share homology with the fragments are  
represented in AAB64821 - AAB64880. The genes and proteins have  
activities dependent on the tissues and cells in which they are  
expressed. Examples of their activities and the activities of their  
agonists and antagonists include; immunosuppressive; antiarthritic;  
antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
cerebroprotective; neurotropic; neuroprotective; antibacterial; vincicide;  
fungicide; and ophthalmological. The secreted proteins, polynucleotides,  
antagonists and agonists may be useful in treating, preventing and  
diagnosing diseases and disorders such as autoimmune diseases e.g.  
rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the  
breast or liver, cardiovascular disorders e.g. cardiac arrest,  
cerebrovascular disorders e.g. cerebral ischaemia, angioneurosis, nervous  
system disorders e.g. Alzheimer's disease, infections caused by bacteria,  
viruses and fungi and ocular disorders e.g. corneal infection. The  
polypeptides can also be used to aid wound healing and epithelial cell  
proliferation, to prevent skin aging due to sunburn, to maintain organs  
before transplantation, for supporting cell culture of primary tissues,  
to regenerate tissues and in chemotaxis. The polypeptides can also be  
used as a food additive or preservative to increase or decrease storage  
capabilities. Included in the invention are sequences AAB64772 and  
AAF33095 - AAF33142 which are used in the isolation and characterisation  
of the nucleotide and protein sequences of the invention.

Sequence 1380 BP; 351 A; 301 C; 270 G; 458 T; 0 other:

Query Match 2.5%; Score 37; DB 22; Length 1380;

Best Local Similarity 100.0%; Pred. No. 5.7e-05;

Matches	37	Conservative	0	Mismatches	0	Indels	0	Gaps	0
Qy	1427	AACTAAAAAAAAAAAAAAAAAAAAAAAAA	1463						
Db	1337	AACTAAAAAAAAAAAAAAAAAAAAAAAAA	1373						

RESULT 33  
AAC59580  
ID AAC59580 standard; CDNA; 1629 BP.  
XX  
XX AAC59580;  
AC  
XX 26-JAN-2001 (first entry)  
DT  
XX  
XX Human secreted protein gene 15 SEQ ID NO:25.  
XX

XX	Human; secreted protein; diagnosis; neuroprotective; cytosstatic;
KM	cardioactive; immunomodulatory; muscular active general; vulnerary;
KM	gastrointestinal; nephrotoxic; antinefctive; gynaecological; and
KM	antibacterial; gene therapy; detection; cancer; chromosome marker;
KM	chromosome identification; neural disorder; immune disorder;
KM	muscular disorder; reproductive disorder; gastrointestinal disorder;
KM	pulmonary disorder; cardiovascular disorder; renal disorder;
KM	proliferative disorder; wound healing; infectious disease; preservative;
XX	food additive; ss.
OS	
XX	Homo sapiens.
PN	
XX	WO200056883-A1.
XX	
PD	
XX	28-SEP-2000.
XX	
PF	16-MAR-2000: 2000MO-US06822.
XX	
XX	23-MAR-1999: 99US-0126054.
PR	10-DEC-1999: 99US-0169916.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Ruben SM, Komatsoulis G;
XX	
DR	WPI: 2000-587666/55.
XX	
XX	P-PSDB; AAB34313.
PT	
PT	Human secreted proteins and gene sequences encoding them, useful for
PT	detecting, preventing, and treating disorders such as cancer,
PT	neurological disorders and immune system disorders -
XX	
XX	Claim 1; Page 347-348; 429pp; English.
XX	

CC The polynucleotide sequences given in AAC59566 to AAC59614 encode the  
CC human secreted proteins given in AAB344299 to AAB344348 to  
CC AAB344437 represent human secreted polypeptide sequences and proteins  
CC homologous to them, which are given in the exemplification of the present  
CC invention. Human secreted proteins have activities based on the tissues  
CC and cells the genes are expressed in. Example of activities include:  
CC neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular  
CC active general; vulnerrary; gastrointestinal; nephrotropic;  
CC antineoplastic; gynaecological; and antibacterial. The polynucleotides  
CC can be used for the detection of various disorders such as cancer,  
CC chromosome identification, as chromosome markers, and for numerous other  
CC diagnostic or research purposes. The secreted proteins can be used to  
CC treat disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders, wound healing, and infectious diseases. The proteins can  
CC also be used as a food additive or preservative to increase or decrease  
CC storage capabilities. AAC59557 to AAC59655 and AAB344298 represent  
CC sequences used in the exemplification of the present invention.  
XX  
XX Sequence 1629 BP; 477 A; 303 C; 312 G; 537 T; 0 other;  
Query Match 2.5%; Score 37; DB 21; Length 1629;

	Best Local Similarity	100.0%	Pred. NO. 5.5e-05,	
	Matches	37;	Conservative	0; Mismatches 0; Indels 0; Gaps 0.
Oy	1427	AACTAAAAAAAAAAAAAAAAAAAAA	1463	
Db	1587	AAGTAAAAAAAAAAAAAAAAAAAAA	1623	

RESULT 34	
AAF63820	
ID	AAF63820 standard; cDNA; 1721 BP.
XX	
AC	AAF63820;
XX	
DT	03-APR-2001 (first entry)
XX	
XX	Human secreted protein gene 32 SEQ ID NO:42.

KW Human; immunosuppressive; antilarval; antitumoric; antiparasitic; neurotropic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KM neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
KM vulnereary; autoimmune disease; hyperproliferative disorder; cancer;  
KM cardiovascular disorder; cerebrovascular disorder; infection;  
KM nervous system disorder; ocular disorder; chemotaxis; food additive;  
KW secreted protein; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200077021-A1.  
XX  
PD 21-DEC-2000.  
XX  
PF 01-JUN-2000; 2000MO-US15135.  
XX  
PR 11-JUN-1999; 99US-0138632.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX {ROSE/} ROSEN C A.  
PI Rosen CA, Ruben SM, Komatsoulis GA;  
XX  
DR WP1: 2001-071257/08.  
XX P-PSDB: AAB75271.  
PT Nucleic acid molecules encoding human secreted proteins, used in  
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
PT Parkinson's diseases -  
XX  
PS Claim 1; Page 457; 530pp; English

This invention relates to polynucleotide sequences AAF653789 - AAF63836 which encode human secreted proteins AAB75280 - AAB75287. Included in the invention are protein sequences AAB75288 - AAB75341 which are fragments of the secreted proteins and amino acid sequences with which these fragments share homology. Examples of the activities of the proteins and polynucleotides and the activities of their agonists and antagonists include, immunosuppressive; antiarthritic; antihematitic; antiproliferative; cytostatic; cardiatic; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; and vulnerary activity. The protein and polynucleotide sequences, their agonists and antagonists may be useful for treating, preventing and diagnosing diseases and disorders such as autoimmune diseases e.g., rheumatoid arthritis, hyperproliferative disorders e.g., neoplasms of the breast or liver, cardiovascular disorders e.g., cardiac arrest, cerebrovascular disorders e.g., cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. Included in the invention are



```
XX New nucleic acid molecules of core 2
PT beta-1,6-N-acetylglucosaminyltransferase useful for providing new
PT compositions for treatment of disorders mediated by the enzyme
PT including cancer, cardiovascular and inflammatory disorders.
XX
PS Claim 4: Page 51-52; 66pp; English.
XX
CC The present sequence encodes a partial human core 2
CC beta-1,6-N-acetylglucosaminyltransferase (core2b G1CNC-r) polypeptide.
CC The polypeptide can be used to treat diseases and disorders, such as
CC cancer, cardiovascular disorders and inflammatory disorders including
CC asthma, rheumatoid arthritis, inflammatory bowel disease,
CC arteriosclerosis, septic shock, adult respiratory distress syndrome
CC (ARDS) and cancer. Various platelet-mediated pathologies such as
CC atherosclerosis and clotting can also be treated. The polypeptides of
CC the invention are predominantly expressed in gastrointestinal tissue
CC (stomach, colon, intestine, testis) and are elevated in cancer.
CC Gastrointestinal disorders that may be prevented or treated include
CC ascites, cholelithiasis, cirrhosis, Crohn's disease, diverticulitis
CC and ulcerative colitis. The antibodies may be used in
CC immuno-histochemical analysis, to detect the novel polypeptide and to
CC localize it to particular cells and tissues and to specific subcellular
CC locations and to quantitate the level of expression.
XX
SQ Sequence 2108 BP; 596 A; 464 C; 509 G; 539 T; 0 other:
Query Match
Best Local Similarity 2.5%; Score 37; DB 21; Length 2108;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2036 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2072
RESULT 37
ID AAA23424 standard; cDNA; 2198 BP.
XX
AC AAA23424;
XX
DT 19-JUN-2000 (first entry)
XX
DE cDNA encoding human secreted protein vb12.1, SEQ ID NO:3.
XX
KW Human; secreted protein; cancer; tumour; cardiovascular disorder;
KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;
KW neurodegenerative disease; asthma; contraceptive; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 63..485
FT /tag= a
FT /product= "Human secreted protein vb12.1"
XX
PD WO200011015-A1.
XX
PD 02-MAR-2000.
XX
PF 24-AUG-1999; 99WO-0519351.
XX
PR 24-AUG-1998; 98US-0097638.
PR 24-AUG-1998; 98US-0097659.
PR 09-SEP-1998; 98US-0099618.
PR 28-SEP-1998; 98US-0102092.
PR 25-NOV-1998; 98US-0109978.
PR 23-DEC-1998; 98US-0113645.
PR 23-DEC-1998; 98US-0113646.
PR 23-AUG-1999; 99US-0379246.
XX
```

```
PA (ALPH-) ALPHAGENE INC.
XX
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P.
XX WPI: 2000-224657/19.
DR P-PSDB; AAY94982.
XX
PT New secreted or transmembrane proteins and polynucleotides encoding
PT them, useful for treating neurodegenerative disorders, autoimmune
PT diseases and cancer.
XX
PS Claim 12: Page 264-265; 357pp; English.
XX
CC The invention relates to 40 human secreted proteins (AAY94981-Y95020),
CC and cDNA sequences encoding them (AAA23423-A23462). The secreted
CC proteins of the invention include those that are thought to be only
CC partially secreted, i.e., transmembrane proteins. The proteins of the
CC invention may exhibit one or more activities selected from the following:
CC cytokine activity; cell proliferation; differentiation; immune
CC modulation; haematopoiesis regulation; tissue growth activity;
CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic
CC and thrombolytic activity; anti-inflammatory activity; and tumour
CC inhibition activity. The proteins may be administered to patients as
CC vaccines, and the nucleotides may be used as part of a gene therapy
CC regime. Diseases or conditions that may be treated using the proteins or
CC nucleotides of the invention include autoimmune diseases; genetic
CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
CC fungal and viral infections, especially HIV; multiple sclerosis;
CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
CC insulin dependent diabetes mellitus; and allergic reactions such as
CC asthma and anaemia. They may also be used for treating wounds, burns,
CC disease, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may
CC additionally be useful as contraceptives. Nucleic acid sequences of the
CC invention may be used in chromosome mapping, and as a source of
CC diagnostic primers and probes. The present sequence represents cDNA
CC encoding one of the 40 proteins of the invention.
XX
SQ Sequence 2198 BP; 618 A; 460 C; 531 G; 589 T; 0 other:
Query Match
Best Local Similarity 2.5%; Score 37; DB 21; Length 2198;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2087 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2123
RESULT 38
ID AA197914 standard; cDNA; 2260 BP.
XX
AC AA197914;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human neuroblastoma expressed polynucleotide SEQ ID NO 3899.
XX
KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS Homo sapiens.
XX
PD WO200166719-A1.
XX
PD 13-SEP-2001.
XX
PF 02-MAR-2001; 2001WO-JP01629.
XX
PR 07-MAR-2000; 2000JP-0159195.
XX
PA (CHIB-) CHIBA PREFECTURE.
```

```
PA (HISM ) HISAMITSU PHARM CO LTD.
XX
XX Nakagawara A:
XX
XX WPI: 2001-565584/63.
DR
XX Nucleic acids originating in gene expressed in human neuroblastoma,
PT useful as probe or primer in diagnosing prognosis of human
PT neuroblastoma, malignancy and susceptibility indicator or tumour marker
PT for anti-cancer agents
XX
XX Claim 1; Page 2935-2936; 2979pp; Japanese.
XX
XX The invention relates to novel genes (AA193926-AA197963) expressed in
CC human neuroblastoma. The nucleic acids are applicable as a probe or
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC susceptibility indicators or tumour markers for anti-cancer agents. The
CC gene information for diagnosing prognosis is related to factors similar
CC to that for N-myc and TrkA genes.
XX
SQ Sequence 2260 BP; 677 A; 503 C; 509 G; 571 T; 0 other;

Query Match          2.5%; Score 37; DB 22; Length 2260;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 2152 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2188

RESULT 39
AA198068
ID AA198068 standard; cDNA; 2260 BP.
XX
XX AA198068;
XX
XX 04-DEC-2001 (first entry)
XX
XX Human neuroblastoma expressed polynucleotide SEQ ID NO 11.
DE
XX Human; neuroblastoma; ss.
KW
XX Homo sapiens.
OS
XX Homo sapiens.
PN
XX WO200166733-A1.
XX
XX 13-SEP-2001.
PD
XX
XX 02-MAR-2001; 2001WO-JP01631.
PF
XX
XX 07-MAR-2000; 2000JP-0159195.
PR
XX 12-MAY-2000; 2000JP-0140387.
XX
XX (CHIB-) CHIBA PREFECTURE.
PA
XX (HISM ) HISAMITSU PHARM CO LTD.
XX
XX Nakagawara A;
PI
XX
XX WPI: 2001-602630/68.
DR
XX
XX Nucleic acids for prognosis of human neuroblastoma comprise nucleic
PT acids expressed by human neuroblastomas -
PT
XX
XX Claim 1; Page 51-53; 159pp; Japanese.
XX
XX The invention relates to nucleic acids (AA198058-AA199161) or their
CC homologues expressed by human neuroblastomas useful for detecting genes
CC expressed by neuroblastoma and for analysing their structure and
CC function. The nucleic acids are useful for the diagnosis and prognosis of
CC neuroblastoma.
XX
XX Sequence 2260 BP; 677 A; 503 C; 509 G; 571 T; 0 other;
```

```
Query Match          2.5%; Score 37; DB 22; Length 2260;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 2152 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2188

RESULT 40
AA293309
ID AA293309 standard; DNA; 2297 BP.
XX
XX AA293309;
XX
XX 04-JUL-2000 (first entry)
XX
XX Sequence encoding mouse Homer-3.
DE
XX
XX Homer; calcium; receptor; immediate early gene; IEG;
KW identification; treatment; glutamate receptor;
KW inositol triphosphate; epilepsy; glutamate toxicity;
KW memory disorder; learning disorder; stroke; schizophrenia;
KW Alzheimer's disease; tissue degeneration; brain development;
KW cardiac disorder; muscular disorder; vascular disorder;
KW neurological disorder; psychiatric disorder; renal disorder;
KW uterine disorder; bronchial disorder; ageing; mouse; ss.
XX
XX Mus musculus.
OS
XX
XX Key Location/Qualifiers
FH 1..1177
FT CDS /*tag= a
FT /product= Homer-3
XX
XX WO200011204-A2.
XX
XX 02-MAR-2000.
PD
XX
XX 18-AUG-1999; 99WO-US18973.
PF
XX
XX 18-AUG-1998; 98US-0097334.
PR
XX 09-JUN-1999; 99US-0138426.
PR
XX 09-JUN-1999; 99US-0138493.
PR
XX 09-JUN-1999; 99US-0138494.
XX
XX (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PA
XX
XX Worley PF, Tu JC, Xiao B, Leahy D, Beneken J, Lanhnan AA;
PI
XX
XX WPI: 2000-246571/21.
DR
XX
XX P-PSDB; AAY83013.
XX
XX Identifying compounds capable of modulating cellular response useful
PT for treating Alzheimer's disease and cardiac disorders, involves
PT incubating compound with cell expressing Homer protein and cell surface
PT receptor
PT
XX
XX Example 1; Page 151; 171pp; English.
XX
XX Homer proteins are the products of neuronal immediate early genes
CC (IEG's). They selectively bind the carboxy termini of certain
CC cell-surface receptors, certain intracellular receptors and binding
CC proteins. Many forms of Homer proteins contain a "coiled-coil"
CC structure in the carboxy terminal domain which mediated homo- and
CC heteromultimerisation between Homer proteins. Homer plays a
CC significant role in mediating receptor-activated calcium mobilisation
CC from intracellular stores. Thus, cells expressing a Homer protein
CC can be used to identify a compound capable of modulating a cellular
CC response mediated by cell surface receptor or intracellular receptor.
CC
XX Compounds identified in this manner which modulate Homer protein
CC activity are useful for treating disorders associated with glutamate
```

CC receptors such as epilepsy, glutamate toxicity, memory disorders,  
CC disorders of learning, stroke, schizophrenia, Alzheimer's disease,  
CC tissue degeneration and disorders of brain development and also for  
CC treating disorders associated with Homer protein activity which  
CC renal, uterine and bronchial tissue disorders and for affecting the  
CC natural aging process. These compounds are also useful for modulating  
CC receptor-mediated calcium mobilization, by exposing a cell to the  
CC compound to modulate calcium mobilization that normally occurs  
CC when the cell is exposed to a ligand, typically an agonist or  
CC antagonist of metabotropic glutamate receptors, or to activate an  
CC intracellular signaling pathway, especially an inositol triphosphate  
CC signaling pathway.

XX Sequence 2297 BP; 534 A; 668 C; 652 G; 443 T; 0 other;

Query Match 2.5%; Score 37; DB 21; Length 2297;  
Best Local Similarity 100.0%; Pred. No. 5.2e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 2257 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2293

RESULT 41  
AAL2416

ID AAL2416 standard; CDNA: 2602 BP.

XX AAL2416;

DT 25-JUL-2000 (first entry)

DE CDNA encoding a human RNA-associated protein.

XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;  
KM immune response; reproductive disorder; acinic keratosis;  
KM atherosclerosis; arteriosclerosis; buritis; cirrhosis; hepatitis;  
KM mixed connective tissue disease; myelofibrosis; primary thrombocythemia;  
KM paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;  
KM trauma; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
FT CDS 260..2110  
FT /\*tag= a  
FT /product= "RNA-associated protein"

XX WO200015799-A2.

XX PD 23-MAR-2000.

XX PF 17-SEP-1999; 99WO-US21688.

XX PR 17-SEP-1998; 98US-0156039.

XX PR 22-SEP-1998; 98US-0158720.

XX PR 04-NOV-1998; 98US-0186815.

XX PR 08-APR-1999; 99US-0128660.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;  
XX PI Hillman JL, Baughn MR, Lal P, Azimzal Y, Yue H, Yang J;  
XX DR WPI: 2000-271437/23.  
XX DR P-PSDB: AAT84444.

XX New polypeptides and polynucleotides, useful for preventing and  
XX treating a disorder associated with increased or decreased expression  
XX of RNA associated proteins -  
XX Claim 9; Page 124-125; 131pp; English.

XX The present sequence encodes a human RNA-associated protein. The  
CC expression of RNA-associated proteins is closely associated with  
CC reproductive tissues, nervous tissues, cell proliferation including  
CC cancer, inflammation and immune responses, and so they may be used  
CC for diagnosis, treatment or prevention of cell proliferative,  
CC immune/inflammatory disorders, and reproductive disorders. Diseases  
CC and disorders which may be treated include actinic keratosis,  
CC atherosclerosis, arteriosclerosis, buritis, cirrhosis, hepatitis,  
CC mixed connective tissue disease, myelofibrosis, paroxysmal nocturnal  
CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia  
CC and cancers, and trauma.

XX Sequence 2602 BP; 732 A; 567 C; 708 G; 595 T; 0 other;

Query Match 2.5%; Score 37; DB 21; Length 2602;  
Best Local Similarity 100.0%; Pred. No. 5.1e-05;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 2549 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2585

RESULT 42  
AAD06812

ID AAD06812 standard; DNA: 2827 BP.

XX AAD06812;

DT 10-AUG-2001 (first entry)

DE Human LP8, a PDGF-related protein encoding DNA.

XX Human; LP8; platelet-derived growth factor-related protein; PDGF  
KM cytoskeletal; vulvular; bone growth; therapy; bone fracture;  
KM prophyllaxis; osteoblast proliferation; osteoporosis; muscle loss;  
KM bone loss; endocrine disorder; arthritis; sarcopenia;  
KM periodontal disease; cartilage differentiation; wound healing; ds.  
XX

XX Homo sapiens.

XX Key Location/Qualifiers  
FT CDS 276..1313  
FT /\*tag= a  
FT /product= "Human LP8, a PDGF-related protein"

XX WO200132197-A2.

XX PD 10-MAY-2001.

XX PF 24-OCT-2000; 2000WO-US26272.

XX PR 02-NOV-1999; 99US-0163056.

XX PR 02-NOV-1999; 99US-0163203.

XX PR 02-NOV-1999; 99US-0163204.

XX PA (ELIL) LILLY & CO ELI.

XX PI Hock JM, Na S, Hammond LJ, Kharitonkov A, Krishnan V, Becker GW;  
XX DR WPI: 2001-316384/33.  
XX DR P-PSDB: AAE02649.

XX New LP8 protein or its fragment for promoting bone growth, treating  
XX bone fractures, increasing or maintaining bone density, and treating  
XX osteoporosis, arthritis, sarcopenia and periodontal disease -  
XX Example 1; Page 57-59; 61pp; English.

XX The present sequence is a DNA encoding human LP8, platelet-derived  
XX growth factor (PDGF)-related protein or its fragment. LP8 or its  
XX fragment is useful for promoting bone growth, treating bone fractures,

CC prophylactically increasing or maintaining bone density in a subject  
 CC having a substantially normal bone density, so as to stimulate  
 CC osteoblast proliferation, treating osteoporosis, muscle or bone loss  
 CC due to malignancy, endocrine disorder, arthritis, sarcopena and  
 CC periodontal disease, preventing cartilage differentiation, and  
 CC wound healing.  
 CC  
 SQ Sequence 2827 BP; 864 A; 522 C; 589 G; 852 T; 0 other;  
 Query Match 2.5%; Score 37; DB 22; Length 2827;  
 Best Local Similarity 100.0%; Pred. No. 5e-05;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 Db 2788 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2824  
 RESULT 43  
 ID AAV40744 standard; cDNA; 2836 BP.  
 AC AAV40744;  
 XX  
 DT 23-SEP-1998 (first entry)  
 DE C. felis esterase, nFE72836, coding sequence.  
 XX  
 DE Esterase; flea; protective immune response; carboxylesterase; arthropod;  
 KW haematophagous ectoparasite infestation; nFE72836; ds.  
 XX  
 OS Ctenocephalides felis.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 99..1889  
 FT /\*tag= a  
 XX  
 PN WO9821324-A1.  
 XX  
 PD 22-MAY-1998.  
 XX  
 PF 10-NOV-1997; 97WO-US20598.  
 XX  
 PR 12-NOV-1996; 96US-0747221.  
 XX  
 PA (HESK-) HESKA CORP.  
 XX  
 PI Brandt KS, Silver GM, Wisniewski N;  
 XX  
 DR WPI: 1998-297929/26.  
 DR P-PSDB; AAM57856.  
 XX  
 PT New nucleic acid encoding carboxyl:esterase(s) from fleas - useful,  
 PT e.g. in vaccines, for preventing infestation by haematophagous  
 PT ectoparasites, particularly on cats and dogs  
 XX  
 PS Claim 1; Page 129-133; 230pp; English.  
 XX  
 CC This sequence encodes the flea esterase protein, nFE72836 (the  
 CC complementary strand is shown in AAV40745), of the invention. When  
 CC administered to animals, the protein induces a protective immune  
 CC (antibody) response against carboxylesterase (CE), so they, or  
 CC compositions containing CE proteins, are used, therapeutically or as  
 CC vaccines, to protect particularly mammals and birds, specifically cats  
 CC and dogs, against haematophagous ectoparasite infestation (HEP).  
 CC arthropods generally, including pests of agricultural crops, trees,  
 CC stored goods etc., also those that are vectors of disease. Fragments of  
 CC the DNA can be used as probes and primers for identification or  
 CC production of nucleic acid. Antibodies against the protein can be used  
 CC for passive immunisation, to screen expression libraries: to isolate the  
 CC protein and to target cytotoxic compounds to HEP. The compounds  
 CC containing CE are effective against both adult and larval stages; they

CC target CEs, including juvenile hormone, that are involved in  
 CC development, metamorphosis, feeding, digestion and reproduction.  
 XX  
 SQ Sequence 2836 BP; 1064 A; 421 C; 473 G; 876 T; 2 other;  
 Query Match 2.5%; Score 37; DB 19; Length 2836;  
 Best Local Similarity 100.0%; Pred. No. 5e-05;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 Db 2780 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2816  
 RESULT 44  
 ID AAV40745/C  
 AC AAV40745;  
 XX  
 DT 23-SEP-1998 (first entry)  
 DE C. felis esterase, nFE72836, coding sequence complementary strand.  
 XX  
 DE Esterase; flea; protective immune response; carboxylesterase; arthropod;  
 KW haematophagous ectoparasite infestation; nFE72836; ds.  
 XX  
 OS Ctenocephalides felis.  
 XX  
 PN WO9821324-A1.  
 XX  
 PD 22-MAY-1998.  
 XX  
 PF 10-NOV-1997; 97WO-US20598.  
 XX  
 PR 12-NOV-1996; 96US-0747221.  
 XX  
 PA (HESK-) HESKA CORP.  
 XX  
 PI Brandt KS, Silver GM, Wisniewski N;  
 XX  
 DR WPI: 1998-297929/26.  
 DR P-PSDB; AAM57856.  
 XX  
 PT New nucleic acid encoding carboxyl:esterase(s) from fleas - useful,  
 PT e.g. in vaccines, for preventing infestation by haematophagous  
 PT ectoparasites, particularly on cats and dogs  
 XX  
 PS Claim 1; Page 136-137; 230pp; English.  
 XX  
 CC This sequence is the complementary strand of the DNA encoding the flea  
 CC esterase protein, nFE72836 (see AAV40744 for coding strand), of the  
 CC invention. When administered to animals, the protein induces a  
 CC protective immune (antibody) response against carboxylesterase (CE), so  
 CC they, or compositions containing CE proteins, are used, therapeutically  
 CC or as vaccines, to protect particularly mammals and birds, specifically  
 CC cats and dogs, against haematophagous ectoparasite infestation (HEP).  
 CC arthropods generally, including pests of agricultural crops, trees,  
 CC stored goods etc., also those that are vectors of disease. Fragments of  
 CC the DNA can be used as probes and primers for identification or  
 CC production of nucleic acid. Antibodies against the protein can be used  
 CC for passive immunisation, to screen expression libraries: to isolate the  
 CC protein and to target cytotoxic compounds to HEP. The compounds  
 CC containing CE are effective against both adult and larval stages; they  
 CC target CEs, including juvenile hormone, that are involved in  
 CC development, metamorphosis, feeding, digestion and reproduction.  
 XX  
 SQ Sequence 2836 BP; 877 A; 473 C; 421 G; 1064 T; 1 other;  
 Query Match 2.5%; Score 37; DB 19; Length 2836;  
 Best Local Similarity 100.0%; Pred. No. 5e-05;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Job time : 261.666 secs

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 ||||||||||||||||||||||||||||||||||  
 DB 56 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 20

## RESULT 45

AAD21170  
 ID AAD21170 standard; CDNA; 2836 BP.

AC AAD21170;

XX 15-JAN-2002 (first entry)

DE Ctenocephalides felis carboxylesterase full length cDNA, nfe72836.

KM Carboxylesterase; haematophagous arthropod infestation; protozoa;  
 KW flea allergic dermatitis; nematode; cestode; trematode; vaccine;  
 KM gene therapy; CE: FAD; nfe72836; ss.

XX Ctenocephalides felis.

XX Key Location/Qualifiers

FT CDS 99..1889

FT /\*tag= a

FT /product= "Carboxylesterase full length protein, PFE7596"

FT /note= "This region is specifically claimed in claim 1

FT as SEQ ID NO:28"

XX US6291222-B1.

XX 18-SEP-2001.

XX 09-JAN-1998; 98US-0005051.

XX 09-JAN-1998; 98US-0005051.

XX (HESK-) HESKA CORP.

XX Sliver GM, Wisniewski N;

XX WPI: 2001-647226/74.

XX P-PSDB; AAE12912.

XX New carboxylesterase nucleic acids and proteins useful for protecting  
 PT animals from hematophagous arthropod infestation, which cause diseases  
 PT (e.g. flea allergic dermatitis) and carry infectious agents (e.g.  
 PT nematodes)

PS Claim 1: Column 103-110; 109pp; English.

XX The invention relates to arthropod such as Ctenocephalides felis  
 CC (flea) carboxylesterase (CE) proteins and nucleic acid molecules  
 CC encoding such proteins. Sequences of the invention are particularly  
 CC useful for protecting animals (e.g. humans, dogs, cattle or zoo animals)  
 CC from hematophagous arthropod infestation which cause diseases such  
 CC as flea allergic dermatitis (FAD) and carry infectious agents (e.g.  
 CC nematodes, cestodes, trematodes or protozoa). They are also useful  
 CC in gene therapy and as vaccines. The present sequence is C. felis  
 CC carboxylesterase full length cDNA referred to as nfe72836.

XX Sequence 2836 BP: 1064 A; 421 C; 473 G; 877 T; 1 other;

XX Query Match 2.5%; Score 37; DB 22; Length 2836;

XX Best Local Similarity 100.0%; Pred. No. 5e-05;

XX Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 ||||||||||||||||||||||||||||||||||  
 DB 2781 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2817

Search completed: November 5, 2002, 13:46:49



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 11:01:37 ; Search time 35.8811 Seconds  
(without alignments)  
10015.359 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463  
Sequence: 1 caccgagaatagctcgcgcc.....aaaaaaaaaaaaaaaa 1463

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

- Issued\_Patents\_NA:\*
- 1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*
  - 2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*
  - 3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*
  - 4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*
  - 5: /cgn2\_6/ptodata/1/ina/PCrtns.COMB.seq:\*
  - 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1463	100.0	1463	US-09-426-557-1	Sequence 1, Appli
2	1360	93.0	1541	US-09-426-557-3	Sequence 3, Appli
3	1267	86.6	1381	US-09-426-557-5	Sequence 3, Appli
4	894	61.1	1478	US-09-426-557-7	Sequence 7, Appli
5	37	2.5	958	US-08-757-046A-5	Sequence 5, Appli
6	37	2.5	958	US-09-447-208-5	Sequence 5, Appli
7	37	2.5	958	US-09-135-988-5	Sequence 5, Appli
8	37	2.5	958	US-09-277-716-5	Sequence 5, Appli
9	37	2.5	958	US-08-597-274A-5	Sequence 5, Appli
10	37	2.5	2836	US-08-747-221B-24	Sequence 24, Appli
11	37	2.5	2836	US-08-747-221B-26	Sequence 24, Appli
12	37	2.5	2836	US-09-005-051-24	Sequence 24, Appli
13	37	2.5	2836	US-09-005-051-26	Sequence 24, Appli
14	36	2.5	2237	US-08-487-135B-1	Sequence 26, Appli
15	36	2.5	2237	US-08-915-672A-1	Sequence 1, Appli
16	35	2.5	2237	US-09-177-909-1	Sequence 1, Appli
17	35	2.4	340	US-08-171-385-27	Sequence 27, Appli
18	35	2.4	340	US-08-361-441B-27	Sequence 27, Appli
19	35	2.4	347	US-08-104-072B-2	Sequence 2, Appli
20	35	2.4	350	US-08-171-385-14	Sequence 2, Appli
21	35	2.4	350	US-08-361-441B-14	Sequence 14, Appli
22	35	2.4	593	US-09-385-982-262	Sequence 16, Appli
23	35	2.4	740	US-08-713-000-8	Sequence 8, Appli
24	35	2.4	740	US-08-975-316-8	Sequence 8, Appli
25	35	2.4	740	US-09-211-710-8	Sequence 8, Appli
26	35	2.4	741	US-08-975-316-8	Sequence 58, Appli
27	35	2.4	1075	US-08-400-006B-6	Sequence 6, Appli

28	35	2.4	1265	US-08-991-789A-169	Sequence 169, App
29	35	2.4	1265	US-09-062-451-169	Sequence 169, App
30	35	2.4	1315	US-09-721-822A-10	Sequence 10, Appli
31	35	2.4	1483	US-09-262-749-1	Sequence 1, Appli
32	35	2.4	1522	US-09-413-574-1	Sequence 1, Appli
33	35	2.4	1551	US-09-461-474-7	Sequence 7, Appli
34	35	2.4	1553	US-09-022-669-1	Sequence 1, Appli
35	35	2.4	1609	US-08-926-808C-2	Sequence 2, Appli
36	35	2.4	1637	US-08-852-824-3	Sequence 3, Appli
37	35	2.4	1639	US-08-333-358-7	Sequence 7, Appli
38	35	2.4	1659	US-08-463-694-7	Sequence 7, Appli
39	35	2.4	1659	US-08-694-501-7	Sequence 7, Appli
40	35	2.4	1728	US-08-985-950-7	Sequence 7, Appli
41	35	2.4	1776	US-08-655-352-10	Sequence 10, Appli
42	35	2.4	1824	US-08-606-505B-1	Sequence 1, Appli
43	35	2.4	1824	US-09-616-990-1	Sequence 1, Appli
44	35	2.4	2399	US-09-385-801-1	Sequence 1, Appli
45	35	2.4	2502	US-09-234-332-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-426-557-1  
: Sequence 1, Application US/09426557  
: Patent No. 6232527  
: GENERAL INFORMATION:  
: APPLICANT: Mahajan, Pramod B.  
: TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
: FILE REFERENCE: 0961  
: CURRENT APPLICATION NUMBER: US/09/426,557  
: CURRENT FILING DATE: 1999-10-22  
: EARLIER APPLICATION NUMBER: 60/112,332  
: EARLIER FILING DATE: 1998-12-15  
: NUMBER OF SEQ ID NOS: 10  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 1  
: LENGTH: 1463  
: TYPE: DNA  
: ORGANISM: Zea mays  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (85)...(1221)  
US-09-426-557-1

Query Match 100.0%; Score 1463; DB 4; Length 1463;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CACGAGATAGCTCGCGCGCGGTTCTTCTGCGGCACCTCCGCGTACAGCCGCCGCCCA	60
DB	1	CACGAGATAGCTCGCGCGCGGTTCTTCTGCGGCACCTCCGCGTACAGCCGCCGCCCA	60
QY	61	CCCCCAGCGCGCGCGCGAGAGATGGGATCAAGGGTTGACGAACGCTGGCGGAC	120
DB	61	CCCCCAGCGCGCGCGCGAGAGATGGGATCAAGGGTTGACGAACGCTGGCGGAC	120
QY	121	AATCGCCCAAGCGCGATGAAGAGACAGAGTTCGAGAGCTACTTCGCGCGCAAAATCGCC	180
DB	121	AATCGCCCAAGCGCGATGAAGAGACAGAGTTCGAGAGCTACTTCGCGCGCAAAATCGCC	180
QY	181	GTCGAGCGCGAGATGACATATACCACTTCTGATATGATGGAAGAGACGATGGA	240
DB	181	GTCGAGCGCGAGATGACATATACCACTTCTGATATGATGGAAGAGACGATGGA	240
QY	241	ACTCTCAAAATGAAGCTGTGAAGTCACTAGTCAATTTGCAAGAAATGTTCAACCGGACA	300
DB	241	ACTCTCAAAATGAAGCTGTGAAGTCACTAGTCAATTTGCAAGAAATGTTCAACCGGACA	300
QY	301	ATAAGATTACTGGAAGCGGAGATCAAGCCAGTTATGTTTGTGATGGCAAGCCTCTGTAT	360
DB	301	ATAAGATTACTGGAAGCGGAGATCAAGCCAGTTATGTTTGTGATGGCAAGCCTCTGTAT	360

Db 301 ATAGATTAATGGAAGCGGAATCAAGCCAGTTATGTTTTGATGGAAGCCTTCAT 360  
QY 361 ATGAACAAACAGAGCTTCTAAAGATCTCAAAAAGATGATGCAAC\*AAAGATCG 420  
Db 361 ATGAAGAACAGAGCTTCTAAAGATCTCAAAAAGATGATGCAACAAATCTG 420  
QY 421 ACTGAGCAGTAGAGTAGAGATAAGATGCGATTTGAAAAATGAGCAAGAGACTGTA 480  
Db 421 ACTGAGCAGTAGAGTAGAGATAAGATGCGATTTGAAAAATGAGCAAGAGACTGTA 480  
QY 481 AAGGTACAAAGGACACACACAGAAATTTAAACGGCTATTAAAGCTTATGGGGTTCCT 540  
Db 481 AAGGTACAAAGGACACACACAGAAATTTAAACGGCTATTAAAGCTTATGGGGTTCCT 540  
QY 541 GTTGTAGAGCAGCTTCTAGAGAGAGAGATGCGACCCCTTGTGCTAAACATAG 600  
Db 541 GTTGTAGAGCAGCTTCTAGAGAGAGAGATGCGACCCCTTGTGCTAAACATAG 600  
QY 601 GTTGTAGAGCAGCTTCTAGAGAGAGATGCGACCCCTTGTGCTAAACATAG 660  
Db 601 GTTGTAGAGCAGCTTCTAGAGAGAGATGCGACCCCTTGTGCTAAACATAG 660  
QY 661 CGTCAATTAATGATGATCAAGATCTGGAAGATCTGATGATGATGATGATGATGATG 720  
Db 661 CGTCAATTAATGATGATCAAGATCTGGAAGATCTGATGATGATGATGATGATGATG 720  
QY 721 GTTTGAGAGAGCTTGAACATGACATGAGACAGGATGATGATGATGATGATGATGATG 780  
Db 721 GTTTGAGAGAGCTTGAACATGACATGAGACAGGATGATGATGATGATGATGATGATG 780  
QY 781 TGTGACTATTTGATGATGATCAAGATCTGGAAGATCTGATGATGATGATGATGATGATG 840  
Db 781 TGTGACTATTTGATGATGATCAAGATCTGGAAGATCTGATGATGATGATGATGATGATG 840  
QY 841 CAACATGAGGCTCATAGAAAGATCTGGAAGATCTGATGATGATGATGATGATGATGATG 900  
Db 841 CAACATGAGGCTCATAGAAAGATCTGGAAGATCTGATGATGATGATGATGATGATGATG 900  
QY 901 GAGAGCTGAGCTTACCAAGAGCTGAGAGCTTGTTCAGAGAGCTTAAATGATGATGATGATG 960  
Db 901 GAGAGCTGAGCTTACCAAGAGCTGAGAGCTTGTTCAGAGAGCTTAAATGATGATGATGATG 960  
QY 961 ATTTCGAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Db 961 ATTTCGAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
QY 1021 GATTAATGTTTCAAGAGAGCTGAGAGCTGAGAGAGCTGATGATGATGATGATGATGATG 1080  
Db 1021 GATTAATGTTTCAAGAGAGCTGAGAGCTGAGAGAGCTGATGATGATGATGATGATGATG 1080  
QY 1081 AATTAATGTTTCAAGAGAGCTGAGAGCTGAGAGAGCTGATGATGATGATGATGATGATG 1140  
Db 1081 AATTAATGTTTCAAGAGAGCTGAGAGCTGAGAGAGCTGATGATGATGATGATGATGATG 1140  
QY 1141 CCGCTAAAGAGAGAGCTGAGAGCTGAGAGAGCTGATGATGATGATGATGATGATGATG 1200  
Db 1141 CCGCTAAAGAGAGAGCTGAGAGCTGAGAGAGCTGATGATGATGATGATGATGATGATG 1200  
QY 1201 AAGGCTGTTGAGAGAGAGCTGAGAGCTGAGAGAGCTGATGATGATGATGATGATGATG 1260  
Db 1201 AAGGCTGTTGAGAGAGAGCTGAGAGCTGAGAGAGCTGATGATGATGATGATGATGATG 1260  
QY 1261 CAGGCTGTTGAGAGAGCTGAGAGCTGAGAGAGCTGATGATGATGATGATGATGATGATG 1320  
Db 1261 CAGGCTGTTGAGAGAGCTGAGAGCTGAGAGAGCTGATGATGATGATGATGATGATGATG 1320  
QY 1321 GGTAAAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
Db 1321 GGTAAAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
QY 1381 TACCAAGTAAACAACTTATGCTGTTTACTTCTTGTGCTTTGATGATGATGATGATGATGATG 1440  
Db 1381 TACCAAGTAAACAACTTATGCTGTTTACTTCTTGTGCTTTGATGATGATGATGATGATGATG 1440

QY 1441 AAAAAAAAAAAAAAAAAAAAAA 1463  
Db 1441 AAAAAAAAAAAAAAAAAAAAAA 1463

RESULT 2  
US-09-426-557-3  
; Sequence 3, Application US/09426557  
; Patent No. 6232527  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses  
; FILE REFERENCE: 0961  
; CURRENT APPLICATION NUMBER: US/09/426,557  
; EARLIER FILING DATE: 1999-10-22  
; EARLIER APPLICATION NUMBER: 60/112,332  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 1541  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (79)...(1215)  
US-09-426-557-3

Query Match 93.08; Score 1360; DB 4; Length 1541;  
Best Local Similarity 99.98; Pred. No. 0;  
Matches 1410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 GCGGTTTCTGCGCAGCTCCGGCTCAGCGCGCCGCCCAACCCGCGCAGCCCGCAGAG 80  
Db 15 GCGGTTTCTGCGCAGCTCCGGCTCAGCGCGCCGCCCAACCCGCGCAGCCCGCAGAG 74  
QY 81 CGAGATGGCGATCAAGGTTTACGAAATCTGCTGGCGAGCATGCGCCCAAGCGATGAA 140  
Db 75 CGAGATGGCGATCAAGGTTTACGAAATCTGCTGGCGAGCATGCGCCCAAGCGATGAA 134  
QY 141 GGAGAGAGATTCAGAGCTACTCGCGCGCAAAATGCGCGTGGAGCGCAGCATGAGCAT 200  
Db 135 GGAGAGAGATTCAGAGCTACTCGCGCGCAAAATGCGCGTGGAGCGCAGCATGAGCAT 194  
QY 201 ATACAGTTCCTGATTTGATGTTGGAAGAGACAGCATGGAATCTCTCAAAATGAAGCTGG 260  
Db 195 ATACAGTTCCTGATTTGATGTTGGAAGAGACAGCATGGAATCTCTCAAAATGAAGCTGG 254  
QY 261 TGAAGTCACTAGCTATTTGCAAGAGATGTTCAACCGGACATTAAGATTACTGGAAGCGGG 320  
Db 255 TGAAGTCACTAGCTATTTGCAAGAGATGTTCAACCGGACATTAAGATTACTGGAAGCGGG 314  
QY 321 AATCAGGCAAGTTATGTTTGTGATGGCAAGCTCTCGATTTGGAAGAACAGAGCTTGC 380  
Db 315 AATCAGGCAAGTTATGTTTGTGATGGCAAGCTCTCGATTTGGAAGAACAGAGCTTGC 374  
QY 381 TAAAGATACCTCAAAAGAGATGATGCAACCAAGATCTGACTGAGGACAGTAGAGTAGG 440  
Db 375 TAAAGATACCTCAAAAGAGATGATGCAACCAAGATCTGACTGAGGACAGTAGAGTAGG 434  
QY 441 AGATTAAGATGCGATTTGAAATTTGAGCAGAGAGACGTTAAAGTTCACAAAGGACACAA 500  
Db 435 AGATTAAGATGCGATTTGAAATTTGAGCAGAGAGACGTTAAAGTTCACAAAGGACACAA 494  
QY 501 CGAAGATTTGAAGCGCTATTAGACTTATGAGGAGTTCTGTTTGAAGGACACTTCTGA 560  
Db 495 CGAAGATTTGAAGCGCTATTAGACTTATGAGGAGTTCTGTTTGAAGGACACTTCTGA 554  
QY 561 AGCAGAGCAGATGTCAGGCCCTTTGATTAAGATTAAGTGTTCCTGTTGCTTCTGAGA 620  
Db 555 AGCAGAGCAGATGTCAGGCCCTTTGATTAAGATTAAGTGTTCCTGTTGCTTCTGAGA 614

```
QY 621 AGATATGACTCCCTTACTTTTGGGGCTCCACGGTTCCTGCTATTATGATCCAAAG 680
    |||
Db 615 AGATATGACTCCCTTACTTTTGGGGCTCCACGGTTCCTGCTATTATGATCCAAAG 674
QY 681 TTCCAAGAAAATACCTGTGATGGAATTTGATGTTGCCAAGTTTGGAGAGCTTGAAT 740
    |||
Db 675 TTCCAAGAAAATACCTGTGATGGAATTTGATGTTGCCAAGTTTGGAGAGCTTGAAT 724
QY 741 CACCATGAGACGCTTATGATTTGTGATCCCTGTGTGATGCTATTTGTATGACAT 800
    |||
Db 735 CACCATGAGACGCTTATGATTTGTGATCCCTGTGTGATGCTATTTGTATGACAT 794
QY 801 CAAAGGTATCGGGGGGCAAAACAGCTCTGAAACTTATTCGCAACATGGGTCATAGAAAG 860
    |||
Db 795 CAAAGGTATCGGGGGGCAAAACAGCTCTGAAACTTATTCGCAACATGGGTCATAGAAAG 854
QY 861 CATCTTGGAGAACTCTTAATTAAGACAGATATCAAAATTCCTGAGGACTGGCTTACCAAGA 920
    |||
Db 855 CATCTTGGAGAACTCTTAATTAAGACAGATATCAAAATTCCTGAGGACTGGCTTACCAAGA 914
QY 921 ACTCGACGCTTGTTCAGAGAGCTTATGATGCATGATGATGATTCGAGCTAAATAGAGAC 980
    |||
Db 915 ACTCGACGCTTGTTCAGAGAGCTTATGATGCATGATGATTCGAGCTAAATAGAGAC 974
QY 981 TGCACCTGATGAGGAGGCTCTCATAAAGTTCCGTGTAAGAAATATGTTTCAACGAAGA 1040
    |||
Db 975 TGCACCTGATGAGGAGGCTCTCATAAAGTTCCGTGTAAGAAATATGTTTCAACGAAGA 1034
QY 1041 TCGGGTGACAAAGGCCATAGAGAGATCAAAATTCGCCAAGATAAATCGTCCGAAGAAAG 1100
    |||
Db 1035 TCGGGTGACAAAGGCCATAGAGAGATCAAAATTCGCCAAGATAAATCGTCCGAAGAAAG 1094
QY 1101 ACTCGAGCTCTTTTCAACCAACTGCCACCATGATGACACCGCTTAAACGGAGAGAGAC 1160
    |||
Db 1095 ACTCGAGCTCTTTTCAACCAACTGCCACCATGATGACACCGCTTAAACGGAGAGAGAC 1154
QY 1161 TTCGGATTAACAGACGACGCTGCGAACCAAGAAAGGCTGTGGAAGAGAGAA 1220
    |||
Db 1155 TTCGGATTAACAGACGACGCTGCGAACCAAGAAAGGCTGTGGAAGAGAGAA 1214
QY 1221 AATATCTTGGAGCTTGTGATGTACACTACGACTACGAAAGCAGCGGTGGCGTATCACTT 1280
    |||
Db 1215 AATATCTTGGAGCTTGTGATGTACACTACGACTACGAAAGCAGCGGTGGCGTATCACTT 1274
QY 1281 CGCTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTGGTAAAGTTGGTGCATGTT 1340
    |||
Db 1275 CGCTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTGGTAAAGTTGGTGCATGTT 1334
QY 1341 TCAAGCTGGGGTAACTTAACTGTTTGAAGAGATTGTGTACCAAGTAAACAAACTTAT 1400
    |||
Db 1335 TCAAGCTGGGGTAACTTAACTGTTTGAAGAGATTGTGTGTACCAAGTAAACAAACTTAT 1394
QY 1401 CGCTGTTTTTACTCTCTGCTCTTGAAGTA 1431
    |||
Db 1395 CGCTGTTTTTACTCTCTGCTCTTGAAGTA 1425

RESULT 3
US-09-426-557-5
; Sequence 5, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 5
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)...(1173)
US-09-426-557-5

Query Match      86.6%; Score 1267; DB 4; Length 1381;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 GCCACAGCCGCGCACACAGGATGGGATCAAGGTTTGACGAAGTGCCTGGCGACAT 123
    |||
Db 16 GCCACAGCCGCGCGCACAGGATGGGATCAAGGTTTGACGAAGTGCCTGGCGACAT 75
QY 124 GCGCCCAAGGCGATGAAGAGCAGAAAGTTTCAGAGCTACTTCGCGCCAAATTCGCCGTC 183
    |||
Db 76 GCGCCCAAGGCGATGAAGAGCAGAAAGTTTCAGAGCTACTTCGCGCCAAATTCGCCGTC 135
QY 184 GAGCCAGCATGACATATACCATTCCTGATTTGTAGTTGGAAGACAGGATGAAACT 243
    |||
Db 136 GAGCCAGCATGACATATACCATTCCTGATTTGTAGTTGGAAGACAGGATGAAACT 195
QY 244 CTCACAAATGAACCTGGTGAAGTCACATGATTCGCAAGGATGTTCAACCGGACAATA 303
    |||
Db 196 CTCACAAATGAACCTGGTGAAGTCACATGATTCGCAAGGATGTTCAACCGGACAATA 255
QY 304 AGATTACTGGAACCGGGAATCAAGCCAGTTTATGTTTGTAGTGGCAAGCCTCCTGATATG 363
    |||
Db 256 AGATTACTGGAACCGGGAATCAAGCCAGTTTATGTTTGTAGTGGCAAGCCTCCTGATATG 315
QY 364 AAGAAACAGAGCTTGTCTAAAGATACCTCAAAAAGAGATGATCAACCAAGATCTGACT 423
    |||
Db 316 AAGAAACAGAGCTTGTCTAAAGATACCTCAAAAAGAGATGATCAACCAAGATCTGACT 375
QY 424 GAGCGATGAGTATGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 483
    |||
Db 376 GAGCGATGAGTATGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 435
QY 484 GTACCAAGGCACACACAGCAAGATTTGTAACGCTATTAAGACTTATGAGGCTTCTGTT 543
    |||
Db 436 GTACCAAGGCACACACAGCAAGATTTGTAACGCTATTAAGACTTATGAGGCTTCTGTT 495
QY 544 GTAGAGGACCTTCTGAAGCAGACGAATGTGCAAGCCTTGTGATTAACGATTAAGTGTG 603
    |||
Db 496 GTAGAGGACCTTCTGAAGCAGACGAATGTGCAAGCCTTGTGATTAACGATTAAGTGTG 555
QY 604 TTCGCTGTTGCTTCAAGAGATATGACCTCCCTACTTTGGGGCTCCACGGTCCCTTGGT 663
    |||
Db 556 TTCGCTGTTGCTTCAAGAGATATGACCTCCCTACTTTGGGGCTCCACGGTCCCTTGGT 615
QY 664 CATTTAATGATCAAGTTCACAGAAATACCTGTATGGAATTTGATGTTGGCAAGT 723
    |||
Db 616 CATTTAATGATCAAGTTCACAGAAATACCTGTATGGAATTTGATGTTGGCAAGT 675
QY 724 TTGGAGAGCTTGAATCTCACATGACCAAGTTCATTTGTCATCCTGTGTGATGT 783
    |||
Db 676 TTGGAGAGCTTGAATCTCACATGACCAAGTTCATTTGTCATCCTGTGTGATGT 735
QY 784 GACTATTTGATAGCATCAAGAGTATCGGGGGGCAACAGCTGTGAAGTATTTGTCATA 843
    |||
Db 736 GACTATTTGATAGCATCAAGAGTATCGGGGGGCAACAGCTGTGAAGTATTTGTCATA 795
QY 844 CATGGTCCATAGAAAGCATCTTGAAGATCTTAATTAAGACAGATATCAAAATTCCTGAG 903
    |||
Db 796 CATGGTCCATAGAAAGCATCTTGAAGATCTTAATTAAGACAGATATCAAAATTCCTGAG 855
QY 904 GACTGGCCTTACCAAGAGCTGACGCTTGTTCAGAGACCTTAATGTCACTTGGATATT 963
    |||
Db 856 GACTGGCCTTACCAAGAGCTGACGCTTGTTCAGAGACCTTAATGTCACTTGGATATT 915
```

QY	964	CGGAGCTAAATAGGACCTGCACCTGATTGAGGAGGGGCTCATTAAGTTTCTCGGTAAAGAT	1023
Db	916	CCTGAGCTTAAATGAGACTGCACCTGATTGAGGAGGGCTCATTAAGTTTCTCGGTAAAGAT	975
QY	1024	AATGGTTTCAACCAAGATCGGGGTGACAAAGGCCATTAGAGAGATCAAAATCTGCCAAGAT	1083
Db	976	AATGGTTTCAACCAAGATCGGGGTGACAAAGGCCATTAGAGAGATCAAAATCTGCCAAGAT	1035
QY	1084	AAATCGTGCAGGAAGACTCGAGTCCCTTTTTCAGGCCAAGTCCACACATCAGCACCG	1143
Db	1036	AAATCGTGCAGGAAGACTCGAGTCCCTTTTTCAGGCCAAGTCCACACATCAGCACCG	1095
QY	1144	CTAAGACGGAAGGAGACTTCGGATTAACAAAGCAAGCGCTGCCAACAAAGAAACAAG	1203
Db	1096	CTAAGACGGAAGGAGACTTCGGATTAACAAAGCAAGCGCTGCCAACAAAGAAACAAG	1155
QY	1204	GCTGTGGAAGAAAGAAATATCTTGATGCTGTGATGTAACAATACGACTACGAAAGCAG	1263
Db	1156	GCTGTGGAAGAAAGAAATATCTTGATGCTGTGATGTAACAATACGACTACGAAAGCAG	1215
QY	1264	CGGTGCGGTATCATCTTGCTAGATTATTAACTCCCTGTTTAACTCAGAGCTTGGT	1323
Db	1216	CGGTGCGGTATCATCTTGCTAGATTATTAACTCCCTGTTTAACTCAGAGCTTGGT	1275
QY	1324	AAAAGTT 1330	
Db	1276	AAAAGTT 1282	

```

RESULT 4
US-09-426-557-7
: Sequence 7, Application US/09426557
: Patent No. 6232527
: GENERAL INFORMATION:
: APPLICANT: Mahajan, Pramod B.
: TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
: TITLE OF INVENTION: Thereof
: FILE REFERENCE: 0961
: CURRENT APPLICATION NUMBER: US/09/426,557
: CURRENT FILING DATE: 1999-10-22
: EARLIER APPLICATION NUMBER: 60/112,332
: EARLIER FILING DATE: 1998-12-15
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 7
: LENGTH: 1478
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (97)...(1233)
US-09-426-557-7

```

Query Match	61.1%;	Score 894;	DB 4;	Length 1478;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 1244;	Conservative	0;	Mismatches 7;	Indels 0;
			Gaps	0;

Accession	Sequence	Position
Y7	CGCGGTTTCCTGGCCACATCCGCGTCAAGCCGCCGCCACCCGCCACGCCGCCGAC	91
Db	CGCGGTTTCCTGGCCACATCCGCGTCAAGCCGCCGCCACCCGCCACGCCGCCGAC	91
Qy	ACGAGATGGGCATCAAGGGTTGTACGAAACGCTGGCGGACAAATGGCCCAAGGCGATGA	139
Db	ACGAGATGGGCATCAAGGGTTGTACGAAACGCTGGCGGACAAATGGCCCAAGGCGATGA	139
Qy	ACGAGATGGGCATCAAGGGTTGTACGAAACGCTGGCGGACAAATGGCCCAAGGCGATGA	151
Db	ACGAGATGGGCATCAAGGGTTGTACGAAACGCTGGCGGACAAATGGCCCAAGGCGATGA	151
Qy	AGGAGCAGAAGCTTCGAGAGCTACTTCGGCCGCAAAATCGCGTGAAGCCAGCATGACA	199
Db	AGGAGCAGAAGCTTCGAGAGCTACTTCGGCCGCAAAATCGCGTGAAGCCAGCATGACA	199
Qy	AGGAGCAGAAGCTTCGAGAGCTACTTCGGCCGCAAAATCGCGTGAAGCCAGCATGACA	211
Db	AGGAGCAGAAGCTTCGAGAGCTACTTCGGCCGCAAAATCGCGTGAAGCCAGCATGACA	211
Qy	TATACCACTTCCTCGATGTTAGTTGGAAGAGCAGCATGGAAACTCTCAAAATGAAGCTG	259
Db	TATACCACTTCCTCGATGTTAGTTGGAAGAGCAGCATGGAAACTCTCAAAATGAAGCTG	259
Qy	TCTACCACTTCCTCGATGTTAGTTGGAAGAGCAGCATGGAAACTCTCAAAATGAAGCTG	271
Db	TCTACCACTTCCTCGATGTTAGTTGGAAGAGCAGCATGGAAACTCTCAAAATGAAGCTG	271

QY	260	GTGAAGCACTACTCACTCAATTTGGCAAGCAAACTTTCAACGGCAACAATAAGATTACTGGAAACGG	319
Db	272	GTGAAGTCACTACTCACTCAATTTGGCAAGCAAACTTTCAACGGCAACAATAAGATTACTGGAAACGG	331
QY	320	GAATCAAGCCAGTTTATGTGTTTGGATGGCAACGCTCTGATATGAAGAACAAGACGTTG	379
Db	332	GAATCAAGCCAGTTTATGTGTTTGGATGGCAACGCTCTGATATGAAGAACAAGACGTTG	391
QY	380	CTAAAGATATCTAAAAAGATGATGCACCAAAAGATCTGACTGAGGCACTAAGAGTAG	439
Db	392	CTAAAGATATCTAAAAAGATGATGCACCAAAAGATCTGACTGAGGCACTAAGAGTAG	451
QY	440	GAGATTAAGATGCGATTGAAAAATTGAGCAACGAGCACTGTAAGGTGCAAGGCAACACA	499
Db	452	GAGATTAAGATGCGATTGAAAAATTGAGCAACGAGCACTGTAAGGTGCAAGGCAACACA	511
QY	500	ACGAAAGTTTGAACGGCTATTAAAGACTTATGGGGGTTCTGTTGTAGAGGACACCTTCTG	559
Db	512	ACGAAAGTTTGAACGGCTATTAAAGACTTATGGGGGTTCTGTTGTAGAGGACACCTTCTG	571
QY	560	AAGCAGACGAGATGTGCAGCCCTTTTGCATTAACGATTAAGTGTTCCTTTCCTTCAG	619
Db	572	AAGCAGACGAGATGTGCAGCCCTTTTGCATTAACGATTAAGTGTTCCTTTCCTTCAG	631
QY	620	AAGTATGAGACTCCCTACTCTTTGGGGGTCACAGGTCCTTGCTGATTAATGATGATCAA	679
Db	632	AAGTATGAGACTCCCTACTCTTTGGGGGTCACAGGTCCTTGCTGATTAATGATGATCAA	691
QY	680	GTTCACAGAAATATACCTGTGATGGAATTTGATGTTGCCAAGGTTTGGAGAGCTTGAC	739
Db	692	GTTCACAGAAATATACCTGTGATGGAATTTGATGTTGCCAAGGTTTGGAGAGCTTGAC	751
QY	740	TCACCATGAGACCAAGTTCATTTGTCATCTCTGTGTGATGTGACTATTTGTATAGCA	799
Db	752	TCACCATGAGACCAAGTTCATTTGTCATCTCTGTGTGATGTGACTATTTGTATAGCA	811
QY	800	TCAAGGATATCGGGGGGCAACAGCCTCAAACTTATGCTCAACATGGGTCATAGAAA	859
Db	812	TCAAGGATATCGGGGGGCAACAGCCTCAAACTTATGCTCAACATGGGTCATAGAAA	871
QY	860	GCACTCTTGAGAAATCTTAATAAAGACAGATATCAAAATTCCTGAGGACGAGCCTTAACCAAG	919
Db	872	GCACTCTTGAGAAATCTTAATAAAGACAGATATCAAAATTCCTGAGGACGAGCCTTAACCAAG	931
QY	920	AAGCTGACGCTGTGTTCAAGGAGCCTTAATGTCAATTTGATATTTCTCTGAGCTAAATGGA	979
Db	932	AAGCTGACGCTGTGTTCAAGGAGCCTTAATGTCAATTTGATATTTCTCTGAGCTAAATGGA	991
QY	980	CTGCACTGATGAGGAGGCTCTCAATAGTTTCTCGTAAACATATAGTGGTTTCAACGAG	1039
Db	992	CTGCACTGATGAGGAGGCTCTCAATAGTTTCTCGTAAACATATAGTGGTTTCAACGAG	1051
QY	1040	ATTCGGGTGACAAAGGCCATGAGAGATCAATAATCTGCCAAGATTAATCTGCGCAAGGAA	1099
Db	1052	ATTCGGGTGACAAAGGCCATGAGAGATCAATAATCTGCGCAAGATTAATCTGCGCAAGGAA	1111
QY	1100	GACTCGAGTCCCTTTTCAAGCCAAGTCCACCAATCAGCACCGCTTAAGCGGAAGGAGA	1159
Db	1112	GACTCGAGTCCCTTTTCAAGCCAAGTCCACCAATCAGCACCGCTTAAGCGGAAGGAGA	1171
QY	1160	CTTGCGGATTAACACAGCAAGGACGCTGCGCAACAAAGAAACAAGGCTGTGTGAAAGAAAGA	1219
Db	1172	CTTGCGGATTAACACAGCAAGGACGCTGCGCAACAAAGAAACAAGGCTGTGTGAAAGAAAGA	1231
QY	1220	AATATCTGTGATGCTTGATGTCACTACCTAGCACTAGCAAGAGAGGCGGTGGC	1270
Db	1232	AATATCTGTGATGCTTGATGTCACTACCTAGCACTAGCAAGAGAGGCGGTGGC	1282
RESULT 5			
US-08-757-046A-5			
; Sequence 5, Application US/08757046A			
; Patent No. 5876995			

RESULT 5  
US-08-757-046A-5  
; Sequence 5, Application US/08757046A  
; Patent No. 5876995

```

1 GENERAL INFORMATION:
2 APPLICANT: Bryan, Bruce
3 TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
4 NUMBER OF SEQUENCES: 14
5 CORRESPONDENCE ADDRESS:
6 ADDRESSEE: Brown, Martin, Haller & McClain
7 STREET: 1660 Union Street
8 CITY: San Diego
9 STATE: CA
10 COUNTRY: USA
11 ZIP: 92101-2926
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Diskette
15 COMPUTER: IBM Compatible
16 OPERATING SYSTEM: DOS
17 SOFTWARE: FastSeq Version 1.5
18
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/757,046A
21 FILING DATE: 11-25-96
22 CLASSIFICATION: 435
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 08/597,274
25 FILING DATE: 02-06-96
26
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Seidman, Stephanie L
29 REGISTRATION NUMBER: 33,779
30 REFERENCE/DOCKET NUMBER: 6680-105B
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: 619-238-0999
33 TELEFAX: 619-238-0062
34
35 TELEX:
36
37 INFORMATION FOR SEQ ID NO: 5:
38
39 SEQUENCE CHARACTERISTICS:
40 LENGTH: 958 base pairs
41 TYPE: nucleic acid
42 STRANDEDNESS: single
43 TOPOLOGY: linear
44
45 MOLECULE TYPE: cDNA
46
47 HYPOTHEetical: NO
48
49 ANTI-SENSE: NO
50
51 FRAGMENT TYPE:
52
53 ORIGINAL SOURCE:
54
55 FEATURE:
56
57 NAME/KEY: Coding Sequence
58 LOCATION: 115...702
59 OTHER INFORMATION: apoaeguorin-encoding gene
60 PUBLICATION INFORMATION:
61 AUTHORS: Inouye et al.
62 JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
63 VOLUME: 82
64 PAGES: 3154-3158
65 DATE: (1985)
66
67 DOCUMENT NUMBER: PATENT NO.: 5,093,240
68
69 US-08-757-046A-5
70
71 Query Match 2.5%; Score 37; DB 2; Length 958;
72 Best Local Similarity 100.0%; Pred. NO. 6.2e-07;
73 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
74
75 QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
76 |
77 |
78 |
79 |
80 |
81 |
82 |
83 |
84 |
85 |
86 |
87 |
88 |
89 |
90 |
91 |
92 |
93 |
94 |
95 |
96 |
97 |
98 |
99 |
100 |
101 |
102 |
103 |
104 |
105 |
106 |
107 |
108 |
109 |
110 |
111 |
112 |
113 |
114 |
115 |
116 |
117 |
118 |
119 |
120 |
121 |
122 |
123 |
124 |
125 |
126 |
127 |
128 |
129 |
130 |
131 |
132 |
133 |
134 |
135 |
136 |
137 |
138 |
139 |
140 |
141 |
142 |
143 |
144 |
145 |
146 |
147 |
148 |
149 |
150 |
151 |
152 |
153 |
154 |
155 |
156 |
157 |
158 |
159 |
160 |
161 |
162 |
163 |
164 |
165 |
166 |
167 |
168 |
169 |
170 |
171 |
172 |
173 |
174 |
175 |
176 |
177 |
178 |
179 |
180 |
181 |
182 |
183 |
184 |
185 |
186 |
187 |
188 |
189 |
190 |
191 |
192 |
193 |
194 |
195 |
196 |
197 |
198 |
199 |
200 |
201 |
202 |
203 |
204 |
205 |
206 |
207 |
208 |
209 |
210 |
211 |
212 |
213 |
214 |
215 |
216 |
217 |
218 |
219 |
220 |
221 |
222 |
223 |
224 |
225 |
226 |
227 |
228 |
229 |
230 |
231 |
232 |
233 |
234 |
235 |
236 |
237 |
238 |
239 |
240 |
241 |
242 |
243 |
244 |
245 |
246 |
247 |
248 |
249 |
250 |
251 |
252 |
253 |
254 |
255 |
256 |
257 |
258 |
259 |
260 |
261 |
262 |
263 |
264 |
265 |
266 |
267 |
268 |
269 |
270 |
271 |
272 |
273 |
274 |
275 |
276 |
277 |
278 |
279 |
280 |
281 |
282 |
283 |
284 |
285 |
286 |
287 |
288 |
289 |
290 |
291 |
292 |
293 |
294 |
295 |
296 |
297 |
298 |
299 |
300 |
301 |
302 |
303 |
304 |
305 |
306 |
307 |
308 |
309 |
310 |
311 |
312 |
313 |
314 |
315 |
316 |
317 |
318 |
319 |
320 |
321 |
322 |
323 |
324 |
325 |
326 |
327 |
328 |
329 |
330 |
331 |
332 |
333 |
334 |
335 |
336 |
337 |
338 |
339 |
340 |
341 |
342 |
343 |
344 |
345 |
346 |
347 |
348 |
349 |
350 |
351 |
352 |
353 |
354 |
355 |
356 |
357 |
358 |
359 |
360 |
361 |
362 |
363 |
364 |
365 |
366 |
367 |
368 |
369 |
370 |
371 |
372 |
373 |
374 |
375 |
376 |
377 |
378 |
379 |
380 |
381 |
382 |
383 |
384 |
385 |
386 |
387 |
388 |
389 |
390 |
391 |
392 |
393 |
394 |
395 |
396 |
397 |
398 |
399 |
400 |
401 |
402 |
403 |
404 |
405 |
406 |
407 |
408 |
409 |
410 |
411 |
412 |
413 |
414 |
415 |
416 |
417 |
418 |
419 |
420 |
421 |
422 |
423 |
424 |
425 |
426 |
427 |
428 |
429 |
430 |
431 |
432 |
433 |
434 |
435 |
436 |
437 |
438 |
439 |
440 |
441 |
442 |
443 |
444 |
445 |
446 |
447 |
448 |
449 |
450 |
451 |
452 |
453 |
454 |
455 |
456 |
457 |
458 |
459 |
460 |
461 |
462 |
463 |
464 |
465 |
466 |
467 |
468 |
469 |
470 |
471 |
472 |
473 |
474 |
475 |
476 |
477 |
478 |
479 |
480 |
481 |
482 |
483 |
484 |
485 |
486 |
487 |
488 |
489 |
490 |
491 |
492 |
493 |
494 |
495 |
496 |
497 |
498 |
499 |
500 |
501 |
502 |
503 |
504 |
505 |
506 |
507 |
508 |
509 |
510 |
511 |
512 |
513 |
514 |
515 |
516 |
517 |
518 |
519 |
520 |
521 |
522 |
523 |
524 |
525 |
526 |
527 |
528 |
529 |
530 |
531 |
532 |
533 |
534 |
535 |
536 |
537 |
538 |
539 |
540 |
541 |
542 |
543 |
544 |
545 |
546 |
547 |
548 |
549 |
550 |
551 |
552 |
553 |
554 |
555 |
556 |
557 |
558 |
559 |
560 |
561 |
562 |
563 |
564 |
565 |
566 |
567 |
568 |
569 |
570 |
571 |
572 |
573 |
574 |
575 |
576 |
577 |
578 |
579 |
580 |
581 |
582 |
583 |
584 |
585 |
586 |
587 |
588 |
589 |
590 |
591 |
592 |
593 |
594 |
595 |
596 |
597 |
598 |
599 |
600 |
601 |
602 |
603 |
604 |
605 |
606 |
607 |
608 |
609 |
610 |
611 |
612 |
613 |
614 |
615 |
616 |
617 |
618 |
619 |
620 |
621 |
622 |
623 |
624 |
625 |
626 |
627 |
628 |
629 |
630 |
631 |
632 |
633 |
634 |
635 |
636 |
637 |
638 |
639 |
640 |
641 |
642 |
643 |
644 |
645 |
646 |
647 |
648 |
649 |
650 |
651 |
652 |
653 |
654 |
655 |
656 |
657 |
658 |
659 |
660 |
661 |
662 |
663 |
664 |
665 |
666 |
667 |
668 |
669 |
670 |
671 |
672 |
673 |
674 |
675 |
676 |
677 |
678 |
679 |
680 |
681 |
682 |
683 |
684 |
685 |
686 |
687 |
688 |
689 |
690 |
691 |
692 |
693 |
694 |
695 |
696 |
697 |
698 |
699 |
700 |
701 |
702 |
703 |
704 |
705 |
706 |
707 |
708 |
709 |
710 |
711 |
712 |
713 |
714 |
715 |
716 |
717 |
718 |
719 |
720 |
721 |
722 |
723 |
724 |
725 |
726 |
727 |
728 |
729 |
730 |
731 |
732 |
733 |
734 |
735 |
736 |
737 |
738 |
739 |
740 |
741 |
742 |
743 |
744 |
745 |
746 |
```

```

1 ADDRESS: Heller Ehrman White & McCauliffe
2 STREET: 4250 Executive Square, 7th Floor
3 CITY: La Jolla
4 STATE: CA
5 COUNTRY: USA
6 ZIP: 92037
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Diskette
9 COMPUTER: IBM Compatible
10 OPERATING SYSTEM: DOS
11 SOFTWARE: FastSeq Version 1.5
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/09/447,208
14 FILING DATE:
15 CLASSIFICATION:
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: 08/757,046
18 FILING DATE: 11-25-96
19 CLASSIFICATION:
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: 08/597,274
22 FILING DATE: 02-06-96
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Seidman, Stephanie L
25 REGISTRATION NUMBER: 33,779
26 REFERENCE/DOCKET NUMBER: 24727-105C
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: 619-450-8400
29 TELEFAX: 619-450-8499
30 TELEX:
31 INFORMATION FOR SEQ ID NO: 5:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 958 base pairs
34 TYPE: nucleic acid
35 STRANDEDNESS: single
36 TOPOLOGY: linear
37 MOLECULE TYPE: cDNA
38 HYPOTHEICAL: NO
39 ANTI-SENSE: NO
40 FRAGMENT TYPE:
41 ORIGINAL SOURCE:
42 FEATURE:
43 NAME/KEY: Coding Sequence
44 LOCATION: 115...702
45 OTHER INFORMATION: apoeaquorin-encoding gene
46 PUBLICATION INFORMATION:
47 PUBLICATION INFORMATION: PATENT NO.: 5,093,240
48 AUTHORS: Inouye et al.
49 JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
50 VOLUME: 82
51 PAGES: 3154-3158
52 DATE: (1985)
53 US-09-447-208-5
54
55 Query Match 2.5%; Score 37; DB 3; Length 958;
56 Best Local Similarity 100.0%; Pred. No. 6.2e-07;
57 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0.0;
58
59 QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
60 |||||||||||||||||||||||||||||||||||
61 Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927
62
63 RESULT 7
64 US-09-135-988-5
65 ; Sequence 5, Application US/09135988
66 ; Patent No. 6152358
67 ; GENERAL INFORMATION:
68 ; APPLICANT: Bryan, Bruce
69 ;

```

;; TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE  
;; NUMBER OF SEQUENCES: 14  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Heller Ehrman White & McCauliffe  
;; STREET: 4250 Executive Square, 7th Floor  
;; CITY: La Jolla  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 92037  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq Version 1.5  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/135,988  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/757,046  
;; FILING DATE: 11-25-96  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/597,274  
;; FILING DATE: 02-06-96  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Seidman, Stephanie L.  
;; REGISTRATION NUMBER: 33,779  
;; REFERENCE/DOCKET NUMBER: 24727-105C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 619-450-8400  
;; TELEFAX: 619-450-8499  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 958 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; HYPOTHEICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE:  
;; ORIGINAL SOURCE:  
;; FEATURE:  
;; NAME/KEY: Coding Sequence  
;; LOCATION: 115...702  
;; OTHER INFORMATION: apoaeguorin-encoding gene  
;; PUBLICATION INFORMATION: PATENT NO.: 5,093,240  
;; AUTHORS: Inouye et al.  
;; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
;; VOLUME: 82  
;; PAGES: 3154-3158  
;; DATE: (1985)  
;; US-09-135-988-5  
Query Match 2.5%; Score 37; DB 3; Length 958;  
Best Local Similarity 100.0%; Pred. No. 6.2e-07;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
DB 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 8  
US-09-277-716-5  
; Sequence 5, Application US/09277716A  
; Patent No. 6232107  
; GENERAL INFORMATION:  
; ATTORNEY/AGENT INFORMATION:  
; APPLICANT: Szent-Gyorgyi, Christopher

;; APPLICANT: PROLUME, LTD.  
;; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE  
;; CURRENT APPLICATION NUMBER: US/09/277,716A  
;; EARLIER FILING DATE: 1998-03-26  
;; EARLIER APPLICATION NUMBER: 60/102,939  
;; EARLIER FILING DATE: 1998-10-01  
;; EARLIER APPLICATION NUMBER: 60/089,367  
;; EARLIER FILING DATE: 1998-06-15  
;; EARLIER APPLICATION NUMBER: 60/079,624  
;; EARLIER FILING DATE: 1998-03-27  
;; NUMBER OF SEQ ID NOS: 32  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 5  
;; LENGTH: 958  
;; TYPE: DNA  
;; ORGANISM: Aequorea (luminescent jellyfish)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (115)..(702)  
;; FEATURE:  
;; OTHER INFORMATION: apoaeguorin-encoding gene  
;; PUBLICATION INFORMATION:  
;; PATENT DOCUMENT NUMBER: 5,093,240  
;; PATENT FILING DATE: 1987-10-08  
;; PUBLICATION DATE: 1992-03-03  
;; PUBLICATION INFORMATION:  
;; AUTHORS: Inouye, S.  
;; TITLE: Cloning and sequence analysis of cDNA for the luminescent protein aequ  
;; JOURNAL: Proc. Natl. Acad. Sci. USA  
;; VOLUME: 82(10)  
;; PAGES: 3154-3158  
;; DATE: 1985-05  
;; US-09-277-716-5  
Query Match 2.5%; Score 37; DB 4; Length 958;  
Best Local Similarity 100.0%; Pred. No. 6.2e-07;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
DB 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 9  
US-08-597-274A-5  
; Sequence 5, Application US/08597274A  
; Patent No. 6247995  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; TITLE OF INVENTION: BIOLUMINESCENT NOVELTY ITEMS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brown, Martin, Haller & McClain  
; STREET: 1660 Union Street  
; CITY: San Diego  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92101-2926  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/597,274A  
; FILING DATE: 02/06/96  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seidman, Stephanie L.  
; REGISTRATION NUMBER: 33,779



```
REFERENCE/DOCKET NUMBER: 6680-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 115...702
OTHER INFORMATION: apoaquorin-encoding gene
PUBLICATION INFORMATION:
DOCUMENT NUMBER: 5,093,240
AUTHORS: Inouye et al.
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 82
PAGES: 3154-3158
DATE: (1985)
US-08-597-274A-5

Query Match      2.5%; Score 37; DB 4; Length 958;
Best Local Similarity 100.0%; Pred. No. 6.2e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      |||
DB 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 10
US-08-747-221B-24
Sequence 24, Application US/08747221B
Patent No. 6063610
GENERAL INFORMATION:
APPLICANT: Silver, Gary W.
APPLICANT: Misniewski, Nancy
TITLE OF INVENTION: No. 6063610el Carboxylesterase Nucleic Acid
TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Wordperfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,221B
FILING DATE: No. 6063610ember 12, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: FC-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 24:
```

```
SEQUENCE CHARACTERISTICS:
LENGTH: 2836 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 99...1889
US-08-747-221B-24

Query Match      2.5%; Score 37; DB 3; Length 2836;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      |||
DB 2781 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2817

RESULT 11
US-08-747-221B-26/c
Sequence 26, Application US/08747221B
Patent No. 6063610
GENERAL INFORMATION:
APPLICANT: Silver, Gary W.
APPLICANT: Misniewski, Nancy
TITLE OF INVENTION: No. 6063610el Carboxylesterase Nucleic Acid
TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Wordperfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,221B
FILING DATE: No. 6063610ember 12, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: FC-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 2836 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-747-221B-26

Query Match      2.5%; Score 37; DB 3; Length 2836;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      |||
DB 56 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 20

RESULT 12
```

US-09-005-051-24  
; Sequence 24, Application US/09005051  
; Patent No. 6291222  
; GENERAL INFORMATION:  
; APPLICANT: Silver, Gary W.  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: No. 6291222el Carboxylesterase Nucleic Acid  
; TITLE OF INVENTION: Molecules, Proteins and Uses Thereof  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carol Talkington Verser, Ph.D.  
; ADDRESSEE: Heska Corporation  
; STREET: 1825 Sharp Point Drive  
; CITY: Fort Collins  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80525  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: Wordperfect for Windows, Version 7.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/005,051  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/747,221  
; FILING DATE: No. 6291222el December 12, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verser, Carol Talkington  
; REGISTRATION NUMBER: 37,459  
; REFERENCE/DOCKET NUMBER: FC-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 970/493-7272  
; TELEFAX: 970/484-9505  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2836 nucleotides  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 99..1889  
; US-09-005-051-24

Query Match 2.5%; Score 37; DB 4; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 5.7e-07;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
Db 2781 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2817

RESULT 13  
US-09-005-051-26/C  
; Sequence 26, Application US/09005051  
; Patent No. 6291222  
; GENERAL INFORMATION:  
; APPLICANT: Silver, Gary W.  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: No. 6291222el Carboxylesterase Nucleic Acid  
; TITLE OF INVENTION: Molecules, Proteins and Uses Thereof  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carol Talkington Verser, Ph.D.  
; ADDRESSEE: Heska Corporation  
; STREET: 1825 Sharp Point Drive  
; CITY: Fort Collins  
; STATE: Colorado

COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Wordperfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/005,051  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/747,221  
FILING DATE: No. 6291222el December 12, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: FC-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2836 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-005-051-26

Query Match 2.5%; Score 37; DB 4; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 5.7e-07;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
Db 56 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 20

RESULT 14  
US-08-487-135B-1  
; Sequence 1, Application US/08487135B  
; Patent No. 5821122  
; GENERAL INFORMATION:  
; APPLICANT: Yannick Gullon; Francine Jotereau;  
; APPLICANT: Thierry Boon-Falleur; Sophie Lucas;  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES PEPTIDES  
; TITLE OF INVENTION: WHICH FORM COMPLEXES WITH MHC MOLECULE HLA-A2  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felle & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,135B  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/487,135  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5821122man D  
; REGISTRATION NUMBER: 30,946

```

; REFERENCE/DOCKET NUMBER: LUD 5388
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2237
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-487-135B-1

Query Match 2.5%; Score 36; DB 1; Length 2237;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2196 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2231

RESULT 15
US-08-915-972A-1
; Sequence 1, Application US/08915972A
; Patent No. 5886145
; GENERAL INFORMATION:
; APPLICANT: Yannick Gulloux; Francine Jotereau;
; APPLICANT: Thierry Boon-Falleur; Sophie Lucas;
; APPLICANT: Vincent Brihard
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
; TITLE OF INVENTION: COMPLEXES WITH MHC MOLECULE HLA-A2 AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,972A
; FILING DATE: August 21, 1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,135
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5886145man D
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5388 - JEL/NDH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2237
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-915-972A-1

Query Match 2.5%; Score 36; DB 2; Length 2237;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2196 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2231
```

```

RESULT 16
US-09-177-909-1
; Sequence 1, Application US/09177909
; Patent No. 5958711
; GENERAL INFORMATION:
; APPLICANT: Yannick Gulloux; Francine Jotereau;
; APPLICANT: Thierry Boon-Falleur; Sophie Lucas;
; APPLICANT: Vincent Brihard
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES PEPTIDES WHICH
; TITLE OF INVENTION: FORM COMPLEXES WITH MHC MOLECULE HLA-A2 AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/177,909
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/915,972
; FILING DATE: August 21, 1997
; APPLICATION NUMBER: 08/487,135
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5958711man D
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5388 - JEL/NDH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2237
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-177-909-1

Query Match 2.5%; Score 36; DB 2; Length 2237;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2196 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2231

RESULT 17
US-08-171-385-27/C
; Sequence 27, Application US/08171385
; Patent No. 5527884
; GENERAL INFORMATION:
; APPLICANT: Mary E. Russell
; APPLICANT: Ulrike Utans
; TITLE OF INVENTION: Mediators of Chronic Allograft
; TITLE OF INVENTION: Rejection
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
```

COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM PS/2 Model 502 or 555X  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/171,385  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 05433/006001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ. ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-171-385-27

Query Match 2.4%; Score 35; DB 1; Length 340;  
Best Local Similarity 100.0%; Pred. No. 4,8e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
Db 39 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5

RESULT 18  
US-08-361-441B-27/c  
Sequence 27, Application US/08361441B  
Patent No. 6077948  
GENERAL INFORMATION:  
APPLICANT: Russell, Mary E.  
APPLICANT: Utans, Ulrike  
TITLE OF INVENTION: MEDIATORS OF CHRONIC ALLOGRAFT REJECTION  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/361,441B  
FILING DATE: 21-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/171,385  
FILING DATE: 21-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 05433/014001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906

TELEX: 200154  
INFORMATION FOR SEQ. ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-361-441B-27

Query Match 2.4%; Score 35; DB 3; Length 340;  
Best Local Similarity 100.0%; Pred. No. 4,8e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
Db 39 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5

RESULT 19  
US-08-104-072B-2  
Sequence 2, Application US/08104072B  
Patent No. 5639948  
GENERAL INFORMATION:  
APPLICANT: Michiels, Frank  
APPLICANT: Morioka, Sinji  
APPLICANT: Scheirlinck, Trees  
APPLICANT: Komari, Toshiko  
TITLE OF INVENTION: Stamen-specific Promoters from Rice  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 No. 563948west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/104,072B  
FILING DATE: 05-AUG-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 9200272  
FILING DATE: 06-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 91403352.7  
FILING DATE: 10-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 91402590.3  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 91400318.1  
FILING DATE: 08-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kowalchuk, Katherine M.  
REGISTRATION NUMBER: 36,848  
REFERENCE/DOCKET NUMBER: 8076.93USWO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 347 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: rice

```

; TISSUE TYPE: anther
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..322
; OTHER INFORMATION: /product="cDNA T23"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 333..347
; OTHER INFORMATION: /product="cloning adaptor
; OTHER INFORMATION: sequence"
US-08-104-072B-2

Query Match          2.4%; Score 35; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1429 GTAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1463
DB 285 GTAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 319

RESULT 20
US-08-171-385-14/c
; Sequence 14, Application US/08171385
; Patent No. 5527884
; GENERAL INFORMATION:
; APPLICANT: Mary E. Russell
; APPLICANT: Ulrike Utans
; TITLE OF INVENTION: Mediators of Chronic Allograft
; TITLE OF INVENTION: Rejection
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (version 5.0)
; SOFTWARE: WordPerfect (version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/171.385
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/006001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-171-385-14

Query Match          2.4%; Score 35; DB 1; Length 350;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1429 GTAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1463
DB 42 GTAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 8
```

```

RESULT 21
US-08-361-441B-14/c
; Sequence 14, Application US/08361441B
; Patent No. 6077948
; GENERAL INFORMATION:
; APPLICANT: Russell, Mary E.
; APPLICANT: Utans, Ulrike
; TITLE OF INVENTION: MEDIATORS OF CHRONIC ALLOGRAFT REJECTION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/361,441B
; FILING DATE: 21-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/171,385
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/014001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-361-441B-14

Query Match          2.4%; Score 35; DB 3; Length 350;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1429 GTAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1463
DB 42 GTAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 8

RESULT 22
US-09-385-982-262/c
; Sequence 262, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCNDA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
```

;; SEQ ID NO 262  
;; LENGTH: 593  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)...(593)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-385-982-262

Query Match 2.4%; Score 35; DB 4; Length 593;  
Best Local Similarity 100.0%; Pred. No. 4.6e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1429 GTAAAAA  
Db 74 GTAAAAA

## RESULT 23

US-08-713-000-8  
; Sequence 8, Application US/08713000  
; Patent No. 5850020  
; GENERAL INFORMATION:  
; APPLICANT: Bloksberg, Leonard N.  
; APPLICANT: Havukkala, Ilkka  
; APPLICANT: Grierson, Alastair  
; TITLE OF INVENTION: MATERIALS AND METHODS FOR THE  
; TITLE OF INVENTION: MODIFICATION OF PLANT LIGNIN CONTENT  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spekman Picard PLLC  
; STREET: 2601 Elliott Avenue, Suite 4185  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/713,000  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sleath, Janet  
; REGISTRATION NUMBER: 37,007  
; REFERENCE/DOCKET NUMBER: 11000.1003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-269-0565  
; TELEFAX: 206-269-0563  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 740 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-713-000-8

Query Match 2.4%; Score 35; DB 2; Length 740;  
Best Local Similarity 100.0%; Pred. No. 4.5e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1429 GTAAAAA  
Db 690 GTAAAAA

RESULT 24  
US-08-975-316-8  
; Sequence 8, Application US/08975316

;; Patent No. 5952486  
;; GENERAL INFORMATION:  
;; APPLICANT: BLOKSBERG, Leonard N., HAVUKKALA, Ilkka  
;; APPLICANT: and GRIERSON, Alastair W.  
;; TITLE OF INVENTION: MATERIALS AND METHODS FOR  
;; TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT  
;; NUMBER OF SEQUENCES: 88  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Law Offices of Ann W. Spekman  
;; STREET: 2601 Elliott Avenue, Suite 4185  
;; CITY: Seattle  
;; STATE: WA  
;; COUNTRY: USA  
;; ZIP: 98121  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/975,316  
;; FILING DATE:  
;; CLASSIFICATION: 800  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/713,000  
;; FILING DATE: September 11, 1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: SLEATH, Janet  
;; REGISTRATION NUMBER: 37,007  
;; REFERENCE/DOCKET NUMBER: 11000/1003c1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 206-269-0565  
;; TELEFAX: 206-269-0563  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 740 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
US-08-975-316-8

Query Match 2.4%; Score 35; DB 2; Length 740;  
Best Local Similarity 100.0%; Pred. No. 4.5e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1429 GTAAAAA  
Db 690 GTAAAAA

RESULT 25  
US-09-211-710-8  
; Sequence 8, Application US/09211710A  
; Patent No. 6204434  
; GENERAL INFORMATION:  
; APPLICANT: Bloksberg, Leonard N.  
; APPLICANT: Havukkala, Ilkka  
; APPLICANT: Grierson, Alastair  
; TITLE OF INVENTION: Materials and Methods for the  
; TITLE OF INVENTION: Modification of Plant Lignin Content  
; FILE REFERENCE: 11000.1003c3  
; CURRENT APPLICATION NUMBER: US/09/211,710A  
; CURRENT FILING DATE: 1998-12-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 740  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-211-710-8

Query Match 2.4%; Score 35; DB 4; Length 740;

Best Local Similarity 100.0%; Pred. No. 4.5e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
DB 690 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 724

RESULT 26  
US-08-975-316-58  
; Sequence 58, Application US/08975316  
; Patent No. 5952486  
; GENERAL INFORMATION:  
; APPLICANT: BLOKSBERG, Leonard N., HAYUKKALA, Ilkka  
; APPLICANT: and GRIERSON, Alastair W.  
; TITLE OF INVENTION: MATERIALS AND METHODS FOR  
; TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Law Offices of Ann W. Speckman  
; STREET: 2601 Elliott Avenue, Suite 4185  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/975.316  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/713,000  
; FILING DATE: September 11, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SLEATH, Janet  
; REGISTRATION NUMBER: 37,007  
; REFERENCE/DOCKET NUMBER: 11000/1003C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-269-0565  
; TELEFAX: 206-269-0563  
; TELEEX:  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 741 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-975-316-58  
Query Match 2.4%; Score 35; DB 2: Length 741;  
Best Local Similarity 100.0%; Pred. No. 4.5e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
DB 690 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 724  
RESULT 27  
US-08-400-006B-6  
; Sequence 6, Application US/08400006B  
; Patent No. 6229065  
; GENERAL INFORMATION:  
; APPLICANT: FREYSSINET, Georges  
; SATILAND, Alain  
; TITLE OF INVENTION: PRODUCTION OF PLANTS RESISTANT TO  
; ATTACKS BY SCLEROTINIA SCLEROTIIFORM BY THE INTRODUCTION OF  
; A GENE ENCODING AN OXALATE OXIDASE  
; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: 1737 King Street, Suite 500  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22314-2756  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/400,006B  
; FILING DATE: 06-Mar-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 91/02874  
; FILING DATE: 05-MAR-1991  
; APPLICATION NUMBER: PCT/FR92/00195  
; FILING DATE: 04-MAR-1992  
; APPLICATION NUMBER: US 07/941,135  
; FILING DATE: 03-DEC-1992  
; APPLICATION NUMBER: 08/207,105  
; FILING DATE: 08-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Teresa Stanek Rea  
; REGISTRATION NUMBER: 30,427  
; REFERENCE/DOCKET NUMBER: 022650-189  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1075  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; OTHER INFORMATION:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-08-400-006B-6  
Query Match 2.4%; Score 35; DB 4: Length 1075;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
DB 1011 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1045  
RESULT 28  
US-08-991-789A-169  
; Sequence 169, Application US/08991789A  
; Patent No. 6225054  
; GENERAL INFORMATION:  
; APPLICANT: Frudakis, Tony N.  
; Smith, John M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; NUMBER OF SEQUENCES: 292  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group  
; STREET: 701 Fifth Avenue, Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,789A
FILING DATE: 11-Dec-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 210121.419C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 169:
SEQUENCE CHARACTERISTICS:
LENGTH: 1265 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-991-789A-169
SEQUENCE DESCRIPTION: SEQ ID NO: 169:
Query Match
Best Local Similarity 100.0%; Score 35; DB 4; Length 1265;
Pred. No. 4.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 1226 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1260
RESULT 29
US-09-062-451-169
Sequence 169, Application US/09062451
Patent No. 6344550
GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
APPLICANT: Smith, John M.
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF BREAST CANCER
NUMBER OF SEQUENCES: 297
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/062,451
FILING DATE: 04-APR-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.419C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 169:
SEQUENCE CHARACTERISTICS:
LENGTH: 1265 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
```

```
TOPOLOGY: linear
US-09-062-451-169
Query Match
Best Local Similarity 100.0%; Score 35; DB 4; Length 1265;
Pred. No. 4.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 1226 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1260
RESULT 30
US-09-721-822A-10
Sequence 10, Application US/09721822A
Patent No. 6306606
GENERAL INFORMATION:
APPLICANT: Michael J. Weber
APPLICANT: Jacqueline Wyatt
APPLICANT: Lex M. Cowsett
TITLE OF INVENTION: ANTISENSE MODULATION OF MP-1 EXPRESSION
FILE REFERENCE: RUS-0142
CURRENT APPLICATION NUMBER: US/09/721,822A
CURRENT FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 135
SEQ ID NO 10
LENGTH: 1315
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (147)...(521)
OTHER INFORMATION:
US-09-721-822A-10
Query Match
Best Local Similarity 100.0%; Score 35; DB 4; Length 1315;
Pred. No. 4.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 1261 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1295
RESULT 31
US-09-262-749-1
Sequence 1, Application US/09262749
Patent No. 6261793
GENERAL INFORMATION:
APPLICANT: Whyte, David
APPLICANT: McGuirk, Marjole
APPLICANT: Nunez-Oliva, Irma
APPLICANT: Hockenderry, Tish
APPLICANT: Pal, James
TITLE OF INVENTION: RAS CONVERTING ENDOPROTEASE (RCE) AND
METHODS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road K-6-1-1990
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: MACOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/262,749
FILING DATE:
CLASSIFICATION:
```



```

ATTORNEY/AGENT INFORMATION:
NAME: Thampoe, Immac J
REGISTRATION NUMBER: 36,322
REFERENCE/DOCKET NUMBER: OC0100505
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 298-5061
TELEFAX: (908) 298-5388
INFORMATION FOR SEO ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1483 base pairs
type: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 14..1000
US-09-262-749-1

```



Best Local Similarity 100.0%; Pred. No. 4.2e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 1618 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1652

RESULT 38

US-08-463-694-7  
; Sequence 7, Application US/08463694  
; Patent No. 5696233

GENERAL INFORMATION:  
; APPLICANT: EVANS Ph.D., RONALD M.  
; APPLICANT: MANGELSDORF Ph.D., DAVID J.  
; APPLICANT: ONG Ms., ESTELITA S.  
; APPLICANT: ORO Ph.D., ANTHONY E.  
; APPLICANT: BORGMAYER Ph.D., UWE K.  
; APPLICANT: GIGUERE Ph.D., VINCENT NMN  
; APPLICANT: YAO Mr., TSO-PANG NMN  
; TITLE OF INVENTION: NOVEL RECEPTORS  
; NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
; STREET: 444 So. Flower St., Suite 2000  
; CITY: Los Angeles  
; STATE: CA  
; COUNTRY: US

ZIP: 90071-2921  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,694  
; FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/761,068  
FILING DATE: 17-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
NAME: Reiter Ph.D., Stephen E.  
; REGISTRATION NUMBER: 31192  
REFERENCE/DOCKET NUMBER: P31 8936  
; TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ. ID NO: 7:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 1659 base pairs  
TYPE: nucleic acid  
; STRANDEDNESS: single  
TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
; CLONE: XR2 (XR2.SEG)  
FEATURE:  
; NAME/KEY: CDS  
LOCATION: 148..1470

US-08-463-694-7

Query Match 2.4%; Score 35; DB 1; Length 1659;

Best Local Similarity 100.0%; Pred. No. 4.2e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 1618 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1652

RESULT 39  
US-08-694-501-7

; Sequence 7, Application US/08694501  
; Patent No. 5710004

GENERAL INFORMATION:  
; APPLICANT: EVANS Ph.D., RONALD M.  
; APPLICANT: MANGELSDORF Ph.D., DAVID J.  
; APPLICANT: ONG Ms., ESTELITA S.  
; APPLICANT: ORO Ph.D., ANTHONY E.  
; APPLICANT: BORGMAYER Ph.D., UWE K.  
; APPLICANT: GIGUERE Ph.D., VINCENT NMN  
; APPLICANT: YAO Mr., TSO-PANG NMN  
; TITLE OF INVENTION: NOVEL RECEPTORS  
; NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
; STREET: 444 So. Flower St., Suite 2000  
; CITY: Los Angeles  
; STATE: CA  
; COUNTRY: US

ZIP: 90071-2921  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/694,501  
; FILING DATE: 07-AUG-1996  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333,358  
FILING DATE:  
; APPLICATION NUMBER: US/07/761,068

FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
NAME: Reiter Ph.D., Stephen E.  
; REGISTRATION NUMBER: 31192  
REFERENCE/DOCKET NUMBER: P31 8936  
; TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ. ID NO: 7:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 1659 base pairs  
TYPE: nucleic acid  
; STRANDEDNESS: single  
TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
; CLONE: XR2 (XR2.SEG)  
FEATURE:  
; NAME/KEY: CDS  
LOCATION: 148..1470

US-08-694-501-7

Query Match 2.4%; Score 35; DB 1; Length 1659;

Best Local Similarity 100.0%; Pred. No. 4.2e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
DB 1618 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1652

RESULT 40  
US-08-985-950-7

; Sequence 7, Application US/08985950  
; Patent No. 6140076

GENERAL INFORMATION:  
; APPLICANT: Adema, Gosse Jan  
; TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;  
; NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DNAX Research Institute

```
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,950
FILING DATE: 05-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,279
FILING DATE: 21-MARCH-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,181
FILING DATE: 16-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,252
FILING DATE: 06-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Chung, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0670K
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1204
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1728 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 69..929
NAME/KEY: mat_peptide
LOCATION: 132..929
US-08-985-950-7

Query Match 2.4%; Score 35; DB 3; Length 1728;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 1674 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1708

RESULT 41
US-08-655-352-10
Sequence 10, Application US/08655352
Patent No. 6077991
GENERAL INFORMATION:
APPLICANT: Bachettira W. Poovaiah, Zhihua Liu,
Shameekumar Patil, Daisuke Takezawa
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
PRODUCTION OF MALE-STERILE PLANTS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klargust Sparkman Campbell Leigh &
ADDRESS: Whinston, LLP
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
STREET: Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: United States of America
```

```
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,352
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,449
FILING DATE: October 14, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Alan E.
REGISTRATION NUMBER: 35,123
REFERENCE/DOCKET NUMBER: 4630-45000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1776 base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: Tobacco Ccank cDNA and deduced amino-acid
FEATURE:
NAME/KEY: protein-coding sequence (not including
stop codon)
NAME/KEY: stop codon
LOCATION: nucleotides 20-1570
US-08-655-352-10

Query Match 2.4%; Score 35; DB 3; Length 1776;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 1720 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1754

RESULT 42
US-08-606-505B-1
Sequence 1, Application US/08606505B
Patent No. 6114601
GENERAL INFORMATION:
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: KIKUCHI, Shigeto
APPLICANT: KIKUCHI, Shigeto
APPLICANT: SHIMADA, Yukihisa
APPLICANT: OHBAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELIA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/4
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606,505B
FILING DATE: 23-MAR-1996
PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER: JP44963/92
; FILING DATE: 02-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Perty, Lawrence S.
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-218-2100
; TELEFAX: 212-218-2100
; INFORMATION FOR SEQ ID NO: 1 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1824 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Petunia hybrida
; STRAIN: Falcon Blue
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116 to 1633
; IDENTIFICATION METHOD: by experiment
;
US-08-606-505B-1

Query Match          2.4%; Score 35; DB 3; Length 1824;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAA...1463
Db 1785 GTAAAAA...1819

RESULT 43
US-09-616-990-1
; Sequence 1, Application US/09616990
; Patent No. 6232109
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; KIYOKAWA, Shigeto
; SHIMADA, Yukihisa
; OHBAYASHI, Masaya
; SHIMADA, Ritsuko
; OKINAKA, Yasushi
; TITLE OF INVENTION: NOVEL PLANT GENES
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112-3801
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
; OPERATING SYSTEM: MS-DOS Ver3.30
; SOFTWARE: PATENT AID Ver1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/616,990
; FILING DATE: 14-Jul-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP44963/92
; FILING DATE: 02-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Perty, Lawrence S.
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-218-2100
; TELEFAX: 212-218-2200
; INFORMATION FOR SEQ ID NO: 1 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1824 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Petunia hybrida
; STRAIN: Falcon Blue
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116 to 1633
; IDENTIFICATION METHOD: by experiment
; SEQUENCE DESCRIPTION: SEQ ID NO: 1
;
US-09-616-990-1

Query Match          2.4%; Score 35; DB 4; Length 1824;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAA...1463
Db 1785 GTAAAAA...1819

RESULT 44
US-09-385-801-1
; Sequence 1, Application US/09385801
; Patent No. 6180850
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Shi, Jinrui
; TITLE OF INVENTION: Maize Ku70 Orthologue and Uses Thereof
; FILE REFERENCE: 0932
; CURRENT APPLICATION NUMBER: US/09/385,801
; EARLIER APPLICATION NUMBER: 60/098,986
; EARLIER FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2399
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)...(2181)
;
US-09-385-801-1

Query Match          2.4%; Score 35; DB 4; Length 2399;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAA...1463
Db 2356 GTAAAAA...2390

RESULT 45
US-09-234-332-1
; Sequence 1, Application US/09234332A
; Patent No. 6087168
; GENERAL INFORMATION:
; APPLICANT: Cedars-Sinai Medical Center
; APPLICANT: Michel F. Levesque, M.D.
; APPLICANT: Thomas Neuman, Ph.D.
; TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
; FILE REFERENCE: P07 41494
; CURRENT APPLICATION NUMBER: US/09/234,332A
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2502
```

; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (0)..(0)  
; OTHER INFORMATION: Neuro D1 gene: Genbank accession D82347  
US-09-234-332-1

Query Match 2.4%; Score 35; DB 3; Length 2502;  
Best Local Similarity 100.0%; Pred. No. 4.1e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
Db 1822 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1856

Search completed: November 5, 2002, 13:48:45  
Job time : 111.881 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 13:04:56 ; Search time 1251.78 Seconds  
(without alignments)  
15774.429 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463  
Sequence: 1 cacgagaatagctcgcgcc.....aaaaaaaaaaaaa 1463

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: em\_estbta:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_estl2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pin:\*  
16: em\_gss\_vrti:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	485	33.2	554	10 BE639421	BE639421 946033A02
C 2	378	25.8	550	10 BE639422	BE639422 946033A02
C 3	350	23.9	586	9 AI881599	AI881599 606068G09
C 4	334	22.8	467	9 BE186786	BE186786 946012C08
C 5	317	21.7	553	9 AW562789	AW562789 660065H06
C 6	295	20.2	470	9 AI861468	AI861468 614014D03
C 7	265	18.1	456	9 AI065689	AI065689 ag91f12.x
C 8	263	18.0	901	10 BC837708	BC837708 707010C02
C 9	249	17.0	474	9 AW559173	AW559173 660065H06
C 10	245	16.7	532	9 AM000375	AM000375 614014D03
C 11	241	16.5	475	9 AW562517	AW562517 660065H06
C 12	225	15.4	376	9 AI065546	AI065546 ag88e02.x
C 13	224	15.3	470	9 AW288831	AW288831 707010F11
C 14	192	13.1	414	9 AW288784	AW288784 707010C02
C 15	187	12.8	225	9 AW562788	AW562788 660065H06
C 16	148	10.1	470	9 AW288831	AW288831 707010F11
C 17	130	8.9	363	9 AW562518	AW562518 660065H06

C 18	103	7.0	232	10 BF727781	BF727781 1000052F1
C 19	93	6.4	553	9 AI834484	AI834484 606068G09
C 20	68	4.6	126	9 AW147048	AW147048 707010C02
C 21	65	4.4	126	9 AW147048	AW147048 707010C02
C 22	57	3.9	225	9 AI947478	AI947478 614047B01
C 23	41	2.8	346	10 BG406288	BG406288 sac29b02.
C 24	41	2.8	468	9 AL121294	AL121294 DKRP762L
C 25	41	2.8	642	10 BI761155	BI761155 603043646
C 26	41	2.8	822	10 BI760815	BI760815 603043846
C 27	40	2.7	202	9 AW307211	AW307211 sf54b11.y
C 28	40	2.7	203	9 AI373782	AI373782 q254906.x
C 29	40	2.7	294	9 AA678838	AA678838 ah05a03.s
C 30	40	2.7	409	10 BG301426	BG301426 kt04b11.y
C 31	40	2.7	411	10 BM378668	BM378668 MEST567-E
C 32	40	2.7	522	9 AA901800	AA901800 NCC1A3T7
C 33	40	2.7	741	10 BF134331	BF134331 601784314
C 34	39	2.7	109	9 AW568974	AW568974 s173909.y
C 35	39	2.7	150	10 BI319195	BI319195 949026H11
C 36	39	2.7	187	10 BF749502	BF749502 IL0-BN042
C 37	39	2.7	201	10 BF749503	BF749503 IL0-BN042
C 38	39	2.7	226	10 BF764516	BF764516 RCI-CS006
C 39	39	2.7	320	9 AW234203	AW234203 sf22d11.y
C 40	39	2.7	365	12 A2953545	A2953545 2M0218A19
C 41	39	2.7	431	9 AA121171	AA121171 z188905.s
C 42	39	2.7	432	9 AA121150	AA121150 z188907.s
C 43	39	2.7	444	10 BM357517	BM357517 N16IV-E8
C 44	39	2.7	448	9 AI362270	AI362270 qy50e11.x
C 45	39	2.7	593	9 AL501532	AL501532 AL501532

#### ALIGNMENTS

RESULT 1  
BE639421/c 554 bp mRNA linear EST 30-AUG-2000  
LOCUS BE639421  
DEFINITION 946033A02.x2 946 - tassal primordium prepared by Schmidt lab zea  
ACCESSION BE639421  
VERSION BE639421.1 GI:9952838  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade: Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 554)  
REFERENCE Walbot,V.  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 946033 row: A column: 02.

#### FEATURES

source

Location/Qualifiers  
1..554  
/organism="Zea mays"  
/cultivar="OH43"  
/db\_xref="taxon:4577"  
/clone\_lib="946 - tassal primordium prepared by Schmidt lab"  
/tissue\_type="tassels"  
/dev\_stage="just after the transition from vegetative to inflorescence development"  
/lab\_host="XLOIR"  
/note="Organ: tassels; Vector: HybridZAP; Site\_1: EcoRI; Site\_2: XhoI; George Chuck dissected immature tassels between imm and 3mm. Sharon Stanfield prepared the cDNA





REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 586)  
Walbot V.  
Maize ESTs from various cDNA libraries sequenced at Stanford University  
Unpublished (1999)  
Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 60608 row: G column: 09.  
Location/Qualifiers

FEATURES  
source

1..586  
/organism="Zea mays"  
/cultivar="Oh1043"  
/db\_xref="taxon:4577"  
/clone\_lib="606 - Ear tissue cDNA library from Schmidt lab"  
/tissue\_type="mixed"  
/dev\_stage="ear length from 0.5 cm - 2.0 cm"  
/lab\_host="XLOLR (Stratagene)"  
/note="Organ: Immature ear; Vector: pBR-CMV; Site\_1: EcoRI; Site\_2: XhoI; Mixed ear tissue cDNA library from Schmidt lab"

BASE COUNT 186 a 132 c 159 g 109 t

ORIGIN

Query Match 23.9%; Score 350; DB 9; Length 586;  
Best Local Similarity 99.3%; Pred. No. 1,3e-69;  
Matches 550; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

20 CGCGGTTCTTGGGCCCTCCGGCTCAGCGCGCCGCCACCGCCGCGCAG 79  
|||||  
33 CGGCGTTCTTGGGCCCTCCGGCTCAGCGCGCCGCCACCGCCGCGCAG 92  
|||||  
80 ACGAGATGGGCATCAAGGTTTGACGAACGCTGGCGGACCAATGGGCCGAGATGA 139  
|||||  
93 ACGAGATGGGCATCAAGGTTTGACGAACGCTGGCGGACCAATGGGCCGAGATGA 152  
|||||  
140 AGGAGCAGAAAGTTCGAGAGCTACTTCGCGCGCAAAATCGCGTGCAGCGCAGATGACA 199  
|||||  
153 AGGAGCAGAAAGTTCGAGAGCTACTTCGCGCGCAAAATCGCGTGCAGCGCAGATGACA 212  
|||||  
200 TATACCAAGTTCGATGTAGTTGGAAGACAGCATGAAACTCTACAAATGAGCTG 259  
|||||  
213 TCTACCAAGTTCGATGTAGTTGGAAGACAGCATGAAACTCTACAAATGAGCTG 272  
|||||  
260 GTGAGATCACTAGCATTTGCAAGGATGTTCAACCGACAATAGATTTCTGGAAGCGG 319  
|||||  
273 GTGAGATCACTAGCATTTGCAAGGATGTTCAACCGACAATAGATTTCTGGAAGCGG 332  
|||||  
320 GAATCAAGCAGATTATGTTTGTATGGAAGCGCTCCTGATATGAAAGAACAGAGCTTG 379  
|||||  
333 GAATCAAGCAGATTATGTTTGTATGGAAGCGCTCCTGATATGAAAGAACAGAGCTTG 392  
|||||  
380 CTAAAGATCTCAAAAAGAGATGATCAACCAAGATCTGACTGAGGCGATGAGGTAG 439  
|||||  
393 CTAAAGATCTCAAAAAGAGATGATCAACCAAGATCTGACTGAGGCGATGAGGTAG 452  
|||||  
440 GAGATAAGATCGATGAAGAAATTTAGCAAGAGACTGTAAAGGTACAAAGGACAACA 499  
|||||  
453 GAGATAAGATCGATGAAGAAATTTAGCAAGAGACTGTAAAGGTACAAAGGACAACA 512  
|||||  
500 ACGAAGATTTAAACGGCTATTAAGACTTATGGGGTTCTGTTGTTGAGAGCACCTTCTG 559  
|||||  
513 ACGAAGATTTAAACGGCTATTAAGACTTATGGGGTTCTGTTGTTGAGAGCACCTTCTG 572  
|||||  
560 AAGCAGAGCAGAA 573

Db 573 AAGCAGAGCAGAA 586  
|||||

RESULT 4  
LOCUS BE186786/c 467 bp mRNA linear EST 22-JUN-2000  
DEFINITION 946012C08.X1 946 - tassal primordium prepared by Schmidt lab Zea mays cDNA, mRNA sequence.  
ACCESSION BE186786  
VERSION BE186786.1 GI:8665970  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 467)  
Walbot V.  
Maize ESTs from various cDNA libraries sequenced at Stanford University  
Unpublished (1999)  
Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 946012 row: C column: 08.  
Location/Qualifiers

FEATURES  
source

1..467  
/organism="Zea mays"  
/cultivar="Oh43"  
/db\_xref="taxon:4577"  
/clone\_lib="946 - tassal primordium prepared by Schmidt lab"  
/tissue\_type="tassels"  
/dev\_stage="just after the transition from vegetative to inflorescence development"  
/lab\_host="XLOLR"  
/note="Organ: tassels; Vector: HybriZAP; Site\_1: EcoRI; Site\_2: XhoI; George Chack dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."

BASE COUNT 108 a 110 c 96 g 153 t

ORIGIN

Query Match 22.8%; Score 334; DB 9; Length 467;  
Best Local Similarity 99.7%; Pred. No. 5.6e-66;  
Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

886 AGATATCAATTTCTGAGAGCTGCGCTTACCAAGAGCTGAGCCTTTGTAAGAGCTT 945  
|||||  
467 AGATATCAATTTCTGAGAGCTGCGCTTACCAAGAGCTGAGCCTTTGTAAGAGCTT 408  
|||||  
946 AATGTCAATTTGATTTCTCGAGACTAAATGAGCTGACCTGATGAGAGGCTCATA 1005  
|||||  
407 AATGTCAATTTGATTTCTCGAGACTAAATGAGCTGACCTGATGAGAGGCTCATA 348  
|||||  
1006 AGTTTCTGTTAAAGATTAATGTTTCAACGAAATCGGCTGACAAAGGCCATAGAGAA 1065  
|||||  
347 AGTTTCTGTTAAAGATTAATGTTTCAATGAAATCGGCTGACAAAGGCCATAGAGAA 288  
|||||  
1066 ATCAAAATCTGCCAAGATTAATGTTTCAACGAAAGAGCTGAGCTCTTTTCAAGCCAAT 1125  
|||||  
287 ATCAAAATCTGCCAAGATTAATGTTTCAACGAAAGAGCTGAGCTCTTTTCAAGCCAAT 228  
|||||  
1126 GCCACGACATCAGCAGCGCTAAACGGAAGAGAGCTTCGATTAACAGCAAGGAGAGCT 1185  
|||||  
227 GCCACGACATCAGCAGCGCTAAACGGAAGAGAGCTTCGATTAACAGCAAGGAGAGCT 168  
|||||

QY 1186 GCGAACAAGAAACAAAGCGTGTGGAGAAAGAAATTAATCTTGATGCTTGATGATCAAA 1245  
|||||  
Db 167 GCGAACAAGAAACAAAGCGTGTGGAGAAAGAAATTAATCTTGATGATGATCAAA 108  
QY 1246 CTAGACTAGAAAGCAGCGGTGGC 1270  
|||||  
Db 107 CTAGACTAGAAAGCAGCGGTGGC 83

RESULT 5  
AM562789 553 bp mRNA linear EST 10-MAR-2000  
LOCUS 660065H06.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,  
ACCESSION AM562789  
VERSION AM562789.1 GI:7216667  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 553)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 660065 row: H column: 06.

FEATURES  
source location/Qualifiers  
1..553  
/organism="Zea mays"  
/cultivar="Ohio43"  
/db\_xref="taxon:4577"  
/clone\_lib="660 - Mixed stages of anther and pollen"  
/tissue\_type="whole premeiotic anthers to pollen shed"  
/dev\_stage="premeiotic anthers to pollen shed"  
/lab\_host="XLOLR"  
/note="Organ: anthers; Vector: Lambda Zap; Site\_1: EcoRI;  
Site\_2: XhoI; Anther and pollen cDNA library  
Directionally sequenced with 5' end at the EcoRI site.  
Created by Amie Franklin."

BASE COUNT 171 a 129 g 142 t  
ORIGIN

Query Match 21.7%; Score 317; DB 9; Length 553;  
Best Local Similarity 99.7%; Pred. No. 3.2e-62;  
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 903 GGACTGGCTTACCACAAGCTGACGCTTGTCGAAGGAGCCTTAATGTCACATTGGATAT 962  
|||||  
Db 4 GAGATGGCCTTACCACAAGCTGACGCTTGTCGAAGGAGCCTTAATGTCACATTGGATAT 63  
QY 963 TCTGAGCTAAATGAGACTGCACCTGATGAGGAGGCTTCATAAGTTTCTGGTAAAGA 1022  
|||||  
Db 64 TCCGAGCTAAATGAGACTGCACCTGATGAGGAGGCTTCATAAGTTTCTGGTAAAGA 123  
QY 1023 TAATGTTTCAAGCAATGGGGGAGCAAGGCGCATAGAAATCAATCTGCCAAGA 1082  
|||||  
Db 124 TAATGTTTCAATGATGGGGGAGCAAGGCGCATAGAAATCAATCTGCCAAGA 183  
QY 1083 TAAATCGTCCCAAGGAAGATCGAGTCTTTTCAAGCCAACTGCCACACATCAGCACC 1142  
|||||  
Db 184 TAAATCGTCCCAAGGAAGATCGAGTCTTTTCAAGCCAACTGCCACACATCAGCACC 243  
QY 1143 GCTTAAACGGAAGAGACTTCGATTAACCAAGCAGGCGTGGCAGACAAAGAAACAA 1202  
|||||

Db 244 GCTAAACGGAAGAGACTTCGATTAACCAAGCAGGCGTGGCAGACAAAGAAACAA 303  
QY 1203 GGCCTGTGGAAGAAAGATATATCTTGATGCTTGATGATGATGATGATGATGATGAT 1262  
|||||  
Db 304 GGCCTGTGGAAGAAAGATATATCTTGATGCTTGATGATGATGATGATGATGATGAT 363  
QY 1263 GCGGTGGC 1270  
|||||  
Db 364 GCGGTGGC 371

RESULT 6  
A1861468/c 470 bp mRNA linear EST 19-JUL-1999  
LOCUS 614014D03.x1 614 - root cDNA library from Walbot Lab Zea mays cDNA,  
DEFINITION mRNA sequence.  
ACCESSION A1861468  
VERSION A1861468.1 GI:5525575  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 470)  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 614014 row: D column: 03.

FEATURES  
source location/Qualifiers  
1..470  
/organism="Zea mays"  
/cultivar="W23"  
/db\_xref="taxon:4577"  
/clone\_lib="614 - root cDNA library from Walbot Lab"  
/tissue\_type="root"  
/dev\_stage="3-4 days old"  
/lab\_host="XLOLR"  
/note="Organ: root; Vector: pBluescriptII SK+; Site\_1:  
EcoRI; Site\_2: XhoI; 3-4 days old root tissue from Walbot  
Lab (LM)"

BASE COUNT 112 a 113 c 93 g 152 t  
ORIGIN

Query Match 20.2%; Score 295; DB 9; Length 470;  
Best Local Similarity 99.7%; Pred. No. 2.9e-57;  
Matches 345; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 995 CCTGATGAGGAGGCTCATATAGTTTCTGGTAAAGATTAATGTTCAACGAAGATCGG 1044  
|||||  
Db 422 CCTGATGAGGAGGCTCATATAGTTTCTGGTAAAGATTAATGTTTCAACGAAGATCGG 363  
QY 1045 GTGACAAGGCCCATAGAGATCAATCTGCCAAGATTAATGTCGCAAGGAAGATC 1104  
|||||  
Db 362 GTGACAAGGCCCATAGAGATCAATCTGCCAAGATTAATGTCGCAAGGAAGATC 303  
QY 1105 GAGTCTTTTTCACGCACTGCCACACATCAGACCGCTAAACGGAAGAGACTTCG 1164  
|||||  
Db 302 GAGTCTTTTTCACGCACTGCCACACATCAGACCGCTAAACGGAAGAGAGACTTCG 243  
QY 1165 GATAAACAAGCAAGGCGTGGCGAACAAGAAACAAAGCTGGTGAAGAAAGAAATTA 1224  
|||||  
Db 242 GATAAACAAGCAAGGCGTGGCGAACAAGAAACAAAGGCTGGTGAAGAAAGAAATTA 183  
QY 1225 TCTTGATGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1284  
|||||

|||||  
Db 182 TCTTGATGCTTATGATGACACTACGACTACGAAGCAGCGTGCGCTGATCACTTGCCT 123  
QY 1285 TAGATTATTAACCTCCGTGTTAACTACAGAGCTTTGGTAAAGTT 1330  
|||||  
Db 122 TAGATTATTAACCTCCGTGTTAACTACAGAGCTTTGGTAAAGTT 77  
RESULT 7  
A1065689/c 456 bp mRNA linear EST 24-JUL-1998  
LOCUS ag91f12.x1 maize inflorescence immature ear library Zea mays cDNA  
DEFINITION clone ag91f12 3', mRNA sequence.  
ACCESSION A1065689  
VERSION A1065689.1 GI:341096  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 456)  
REFERENCE Schutz, K., de la Bastide, M., Gnoj, L., Habermann, K., Huang, E.N.,  
Parnell, L.D., Dedhia, N., Martienssen, R. and McCombie, W.R.  
Expressed sequence tags from Zea mays  
Unpublished (1998)  
Contact: W. Richard McCombie  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org  
Plate: ag91 row: f column: 12  
Seq primer: MJ3 forward universal -21  
High quality sequence stop: 456.  
FEATURES  
source location/Qualifiers  
1. 456  
/organism="Zea mays"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ag91f12"  
/clone\_1ib="maize inflorescence immature ear library"  
/sex="female"  
/tissue\_type="Immature ear"  
/note="Vector: pBLUESCRIPT SK+ (X52325); Site\_1: XhoI;  
Site\_2: EcoRI; This library is described in Schmidt, Hake,  
et al., (1993) Plant Cell 5:729-737. cDNAs are  
directionally cloned into the XhoI and EcoRI sites; XhoI  
is near the polyA tail. Most reads from this library are  
3' in direction. Additional information on this library as  
well as ftp access to all sequences can be found at  
http://www.cshl.org/maizegenome"  
BASE COUNT 112 a 104 c 86 g 152 t 2 others  
ORIGIN  
Query Match 18.1%; Score 265; DB 9; Length 456;  
Best Local Similarity 100.0%; Pred. No. 1.5e-50;  
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|||||  
Db 185 TGCATGCTTATGATGACACTACGACTACGAAGCAGCGTGCGCTGATCACTTGCCTAG 126  
QY 1288 ATTAATTAACTCCCTGTTTAACTC 1312  
|||||  
Db 125 ATTAATTAACTCCCTGTTTAACTC 101  
RESULT 8  
BC837708 901 bp mRNA linear EST 25-MAY-2001  
LOCUS Zm10\_01f08.A Zm10\_AAFc\_ECORC\_Fusarium\_graminearum\_corn\_silk Zea  
DEFINITION mays cDNA clone Zm10\_01f08, mRNA sequence.  
ACCESSION BC837708  
VERSION BC837708.1 GI:14204031  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 901)  
REFERENCE Harris, L.J., Balcerzak, M., Allard, S., Saparno, A., Couroux, P., De  
Mours, A., Hattori, J.I., Ouellet, T., Robert, L.S., Singh, J.A., Sprott  
D. and Tinker, N.A.  
Expressed Sequence Tags from Maize Silk Six Hours After Silk  
Channel Inoculation with Fusarium graminearum  
Unpublished (2001)  
Contact: Harris, Linda J.  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-Food Canada  
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
CANADA  
Tel: (613) 759-1314  
Fax: (613) 759-6566  
Email: harris1j@em.agr.ca.  
FEATURES  
source location/Qualifiers  
1. 901  
/organism="Zea mays"  
/cultivar="CO388"  
/db\_xref="taxon:4577"  
/clone="Zm10\_01f08"  
/clone\_1ib="Zm10\_AAFc\_ECORC\_Fusarium\_graminearum\_corn\_silk"  
/tissue\_type="Silk"  
/dev\_stage="4-5 days post-silk emergence"  
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site\_1: EcoRI;  
Site\_2: XhoI; Field-grown corn was silk channel-inoculated  
in the morning (~10 am) with 1 ml of a macroconidial  
suspension (500,000 spores/ml) of Fusarium graminearum and  
silk channels were collected and immediately frozen in  
liquid nitrogen 6 hours later. RNA was extracted from  
silk tissue between 1 cm below and above the inoculation  
point in the silk channel, RNA from five silk channels was  
pooled."  
BASE COUNT 267 a 174 c 208 g 249 t 3 others  
ORIGIN  
Query Match 18.0%; Score 263; DB 10; Length 901;  
Best Local Similarity 99.3%; Pred. No. 2.8e-50;  
Matches 603; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

OY	856	GAAGCATTCTGGAGAACTCTTAATAAAGACAGATATCAAAATTC-C-TGAAGCATGGCCTTA	914
Db	252	GAAGCATTCTGGAGAACTCTTAATAAAGACAGATATCAAAATTC	311
OY	915	CC-AAGAGCTCGACGCTTGTCAAGAAGCCTTAATGTCACATTTGATATCTCTGAAGCTAA	973
Db	312	CCAAAGAAGCTCGACGCTTGTCAAGAAGCCTTAATGTCACATTTGATATCTCTGAAGCTAA	371
OY	974	AATGACATGACCCCTATGTAGAGGGCTTCATTAAGTTCTCGTAAAGTAATATGATTTCA	1033
Db	372	AATGACATGACCCCTATGTAGAGGGCTTCATTAAGTTCTCGTAAAGTAATATGATTTCA	431
OY	1034	ACGAGATCGGGGTGACAAAGGCCATAGAGAAAGTAAATCTCCCAAGATAAATGCTGCG	1093
Db	432	ATGAAGATCGGGGTGACAAAGGCCATAGAGAAAGTAAATCTCCCAAGATAAATGCTGCG	491
OY	1094	AAGGAAGACTGAGTCTCTTTTTCAAAGCCAACTGCCCACCATCAGACCCGCTAAACGGA	1153
Db	492	AAGGAAGACTGAGTCTCTTTTTCAAAGCCAACTGCCCACCATCAGACCCGCTAAACGGA	551
OY	1154	AGGAGACTTCGGATAAACAACCAAGCGCAGCTGCCAACAGAAAGAAACAAAGCTGTGGGAA	1213
Db	552	AGGAGACTTCGGATAAACAACCAAGCGCAGCTGCCAACAGAAAGAAACAAAGCTGTGGGAA	611
OY	1214	AGAGCAAAATTAATCTTGATGCTTGTATGTATCAACTACGACTACGAAAGCAGCGGTGGCGTG	1273
Db	612	AGAGCAAAATTAATCTTGATGCTTGTATGTATCAACTACGACTACGAAAGCAGCGGTGGCGTG	671
OY	1274	ATTCACCTT 1280	
Db	672	ATTCACCTT 678	

RESULT	9
LOCUS	AM559173/c
DEFINITION	AM559173 474 bp mRNA linear EST 07-MAR-2000 660065H06.xl 660 - Mixed stages of anther and pollen Zea mays cDNA.
ACCESSION	AM559173
VERSION	AM559173
KEYWORDS	AM559173.1 GI:7204640
SOURCE	EST.
ORGANISM	Zea mays. Zea mays. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 474) Walbot.V. Maize ESTs from various cDNA libraries sequenced at Stanford University Unpublished (1999) Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 660065 row: H column: 06. Location/Qualifiers I..474
JOURNAL COMMENT	
REFERENCE	
AUTHORS	
TITLE	
FEATURES	
source	

```

/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone.lib="660 - Mixed stages of anther and pollen"
/tissue.type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOUP"
/note="Organ: anthers; Vector: Lambda Zap; Site:1: EcoRI
Site:2: XhoI; Anther and pollen cDNA library.
directionally sequenced with 5' end at the EcoRI site,
created by Amie Franklin."

```

BASE COUNT	113 a	114 c	92 g	154 t	1 others
ORIGIN					
Query Match			17.0%:	Score 249:	DB 9:
Best Local Similarity			99.4%:	Pred. No. 5.6e-47:	
Matches 349:	Conservative	0:	Mismatches	2:	Indels 0:
					Gaps 0:
Qy	920	AAGCTCGACGCTGTGTTTCGAAGGAGCCTATGTCACATTGGATATTCCTAGCGTAAAGATGGA	979		
Db	474	AAGCTCGACGCTGTGTTTCGAAGGAGCCTATGTCACATTGGATATTCCTAGCGTAAAGATGGA	415		
Qy	980	CTGCACCTGATGAGGAGGCGTCTCAATAGTTCTCGTAAAGATATGTTTCAACGAG	1039		
Db	414	CTGCACCTGATGAGGAGGCGTCTCAATAGTTCTCGTAAAGATATGTTTCAATGAG	355		
Qy	1040	ATCGGCGTCAAAAGCCATAGAGAATCAATCTGCCAAGATTAATCGTCCGAAGAA	1099		
Db	354	ATCGGCGTCAAAAGCCATAGAGAATCAATCTGCCAAGATTAATCGTCCGAAGAA	295		
Qy	1100	GACTCGAGTCTTTTTCGAAGCCACTGCCACCACATCGACACCGCTAAAGCGAAGGAGA	1159		
Db	294	GACTCGAGTCTTTTTCGAAGCCAACTGCCACCACATCGACACCGCTAAAGCGAAGGAGA	235		
Qy	1160	CTTCGGATTAACAACGACGAGCTGGAAACAATAAACAAGGCTGTGGAAAGAAGA	1219		
Db	234	CTTCGGATTAACAACGACGAGCTGGAAACAATAAACAAGGCTGTGGAAAGAAGA	175		
Qy	1220	AATATCTTGGATGCTTGATGTACACTACAGACTTCGAAGACAGGCGTGGC	1270		
Db	174	AATATCTTGGATGCTTGATGTACACTACAGACTTCGAAGACAGGCGTGGC	124		

FEATURES	Source
RESULT 10	
LOCUS	AM000375
DEFINITION	AM000375 532 bp mRNA linear EST 08-SEP-1999
ACCESSION	614014D03.y1 614 - root cDNA library from walbot Lab Zea mays cDNA.
VERSION	AM000375
KEYWORDS	AM000375.1 GI:5847296
SOURCE	EST.
ORGANISM	Zea mays.
REFERENCE	Zea mays.
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
JOURNAL	clade; Panicoideae; Andropogoneae; Zea.
COMMENT	1 (bases 1 to 532)
	Walbot,V.
	Maize ESTs from various cDNA libraries sequenced at Stanford
	University
	Unpublished (1999)
	Contact: Walbot V
	Department of Biological Sciences
	Stanford University
	855 California Ave, Palo Alto, CA 94304, USA
	Tel: 650 723 2227
	Fax: 650 725 8221
	Email: walbot@stanford.edu
	Plate: 614014 row: D column: 03.
	Location/Qualifiers
	I..532

```

/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/collection_lib="614 - root cDNA library from Walbot Lab"
/tissue_type="root"
/day_stage="3-4 days old"
/lab_host="XLOLR"
/notes="Organ: root; Vector: pBluescriptII SK+; Site_1:
Ecot1; Site_2: Xho1; 3-4 days old root tissue from Walbot
Lab (LM)"
BASE COUNT      173 a      129 g      127 t
ORIGIN

```

Query Match	16.7%:	Score 245:	DB 9:	Length 532:
Best Local Similarity	100.0%:	Pred. No. 4.1e-46:		
Matches 245:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
QY	953	CATTGATATTCCTAGGTAAATAGTGCACCGCTGATGAGAGGGCTCATAGTTTCC	1012	
Db	288	CATTGGATATTCCTAGGTAAATAGTGCACCGCTGATGAGAGGGCTCATAGTTTCC	347	
QY	1013	TGTTAAAGATTAATGTTTCAACGAGATCGGGTGACAAAGGCCATAGAGATCAAT	1072	
Db	348	TGTTAAAGATTAATGTTTCAACGAGATCGGGTGACAAAGGCCATAGAGATCAAT	407	
QY	1073	CTGCCAAGAATTAATGTCGCGCAAGAAAGTCTCTTTTCAAGCCAACTGCCACCA	1132	
Db	408	CTGCCAAGAATTAATGTCGCGCAAGAAAGTCTCTTTTCAAGCCAACTGCCACCA	467	
QY	1133	CATCGACACCGCTAAAGCAAGAGACTGTGGATTAACAACGAAGGAGGTGGACCA	1192	
Db	468	CATCGACACCGCTAAAGCAAGAGACTGTGGATTAACAACGAAGGAGGTGGACCA	527	
QY	1193	AGAAA 1197		
Db	528	AGAAA 532		
RESULT 11				
LOCUS	AW562517/c	475 bp	mRNA	linear
DEFINITION	660065H06.X2 660 - Mixed stages of anther and pollen Zea mays cDNA.			
ACCESSION	AW562517			
VERSION	AW562517.1	GI:7216395		
KEYWORDS	EST.			
SOURCE	Zea mays.			
ORGANISM	Zea mays.			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.			
AUTHORS	1 (bases 1 to 475)			
TITLE	Walbot.V.			
JOURNAL	Maize ESTs from various cDNA libraries sequenced at Stanford University			
COMMENT	Unpublished (1999) Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 660065 row: H column: 06.			
FEATURES	Source			
	location/Qualifiers			
	1..475			
	/organism="Zea mays"			
	/cultivar="Ohio43"			
	/db_xref="taxon:4577"			
	/clone_lib="660 - Mixed stages of anther and pollen"			
	/tissue_type="whole premeiotic anthers to pollen shed"			
	/dev_stage="premeiotic anthers to pollen shed"			
	/lab_host="X10LR"			
	/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI; Site_2: XhoI; Anther and pollen cDNA library. directionally sequenced with 5' end at the EcoRI site. Created by Ame Franklin."			
	BASE COUNT	124 a	110 c	94 g
	ORIGIN		147 t	
Query Match	16.5%:	Score 241:	DB 9:	Length 475:
Best Local Similarity	99.7%:	Pred. No. 3.5e-45:		
Matches 291:	Conservative 0:	Mismatches 1:	Indels 0:	Gaps 0:
QY	979	ACTGCACCTGATGAGAGGCTCATAGTTTCCGTGTAAGAAAGATTAATGTTTCAACGAA	1038	

Db	475	ACTGCACCTGATGAGAGGGCTTCATTAAGTTTCCGTGTAATAAGATATGTTTCAATGA	416
OY	1039	GATCGCGTGACAAGAAGCCATAGAGAAGATCAAATCTGCCAAGTAATAATCGTCCAAAGA	1098
Db	415	GATCGGGTGACAAGAAGCCCATAGAGAAGATCAATATGTCGAAGATAAATCGTCCAAAGA	356
OY	1099	AGACTCGAGTCTTTTTCAGGCCCAACTGCCACCACATCAGCACCGCTTAAAACGGAAGAG	1158
Db	355	AGACTCGAGTCTTTTTCAGGCCCAACTGCCACCACATCAGCACCGCTTAAAACGGAAGAG	296
OY	1159	ACTTCGGATTAAACAGAACGAGCGAGCTCGCAACGAAGAACCAAGAGCTGTGGAAAGAG	1218
Db	295	ACTTCGGATTAAACAGAACGAGCGAGCTCGCAACGAAGAACCAAGAGCTGTGGAAAGAG	236
OY	1219	AAATATATCTGGATCTGTATGTACAACTAGACACAGCAAGAGCGGTGGC	1270
Db	235	AAATATATCTGGATCTGTATGTACAACTAGACACAGCAAGAGCGGTGGC	184
RESULT	12		
LOCUS	A1065546/c	376 bp	mRNA Linear EST 24-JUL-1998
DEFINITION	ag88e02.xl maize inflorescence immature ear library Zea mays cDNA		
ACCESSION	A1065546		
VERSION	A1065546.1	GI:3340953	
KEYWORDS	EST.		
SOURCE	Zea mays.		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 376) Schutz,K., de la Bastide,M., Gnoj,L., Habermann,K., Huang,E.N., Parnell,L.D., Dedhia,N., Martienssen,R. and McCombie,W.R. Expressed sequence tags from Z. mays Unpublished (1998) Contact: W. Richard McCombie Lita Annenberg Hazen Genome Sequencing Center Cold Spring Harbor Laboratory PO Box 100, Cold Spring Harbor, NY 11724, USA Tel.: 516 367 8884 Fax: 516 367 8874 Email: mccombie@csnl.org Plate: ag88 row: e column: 02 Seq primer: M13 forward universal -21 High quality sequence stop: 376. Location/Qualifiers 1..376 /organism="Zea mays" /cultivar="B73" /db_xref="taxon:4577" /clone="ag88e02" /feature="ag88e02" /feature="immature ear library" /sex="female" /tissue-type="immature ear" /note="Vector: pBLUESCRIPT SK+ (X52325); Site_1: XhoI, Site_2: EcoRI; This library is described in Schmidt, Hake, et al., (1993) Plant Cell 5:729-737. cDNAs are directionally cloned into the XhoI and EcoRI sites; XhoI is near the polyA tail. Most reads from this library are 3' in direction. Additional information on this library as well as ftp access to all sequences can be found at <a href="http://www.csl.lsu.edu/malegenome">http://www.csl.lsu.edu/malegenome</a> "		
FEATURES	source		
BASE COUNT	79 a	99 c	65 g 131 t 2 others
ORIGIN			
Query Match	15.4%	Score 225;	DB 9; Length 376;
Best Local Similarity	99.6%;	Pred. No. 1.5e-41;	
Matches 273;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
OY	1016	TAAAGATTAAGCTTTCAACGAGTAGTCGGTGACAAAGCCATAGAGAAATCAATCTG	1075

Dd	369	ATAATGCTTACAGCAAGACTGGGTGGACAAAGGCCATAGAGAAGTCAAATCTGCCACA	310
Oy	1082	ATAAATCGTCGCAGGAAGACTCGAGTCCCTTTTCAAGCCAATCGCACCATCAGCAGC	1141
Dd	309	ATAAATCGTCGCAGGAAGACTCGAGTCCCTTTTCAAGCCAATCGCACCATCAGCAGC	250
Oy	1142	CGCTAAAACGAGAGACTTGGAATAAACAGCAAGAGCAGCTCGCAACAAGAAACAA	1201
Dd	249	CCCTAAAACGAGAGACTTGGAATAAACAGCAAGAGCAGCTCGCAACAAGAAACAA	190
Oy	1202	AGCGTGGTGAAGAAGAATATACCTGGATCCTT	1236
Dd	189	AGCGTGGTGAAGAAGAATATATCTGGATCCTT	155
<hr/>			
RESULT 14			
LOCUS	AM288784	414 bp	mRNA linear EST 16-JAN-2000
DEFINITION	707010C02.x5 707 - Mixed adult tissues from walbot lab (Sk) Zea		
ACCESSION	mays cDNA, mRNA sequence.		
VERSION	AM288784		
KEYWORDS	AM288784.1 GI:6695706		
SOURCE	EST.		
ORGANISM	Zea mays.		
AUTHORS	Zea mays.		
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 414) Walbot.V. Maize ESTs from various cDNA libraries sequenced at Stanford University		
JOURNAL COMMENT	Unpublished (1999) Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 707010 row: C column: 02.		
FEATURES	Location/Qualifiers		
Source	1..414		
	/organism="Zea mays"		
	/cultivar="W23"		
	/db_xref="taxon:4577"		
	/clone_lib="707 - Mixed adult tissues from Walbot Lab (SK)"		
	/tissue_type="tassel, kernel, silk, husk, root, leaf"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/note="Organ: tassel, kernel, silk, husk, root, leaf; vector: pGDH10; site_1: EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned."		
BASE COUNT	120 a 82 c 101 g 109 t 2 others		
ORIGIN			
Query Match	13.1%;	Score 192;	DB 9; Length 414;
Best Local Similarity	99.0%;	Pred. No. 3.4e-34;	
Matches 392;	Conservative 0;	Mismatches 4;	Indels 0; Gaps 0;
Oy	649	CCAGGTCCTTCGCAATTATGATGATCAATTCACGAANAATACCCTGATGAATTT	708
Dd	6	CCAGGTCCTTCGCAATTATGATGATCAATTCACGAANAATACCCTGATGAATTT	65
Oy	709	GATGTTGCCAAGTTTTGGAGAGCTTGAACTCACCATGAGACCACTTCATTGATTTGTGC	768
Dd	66	CATGTTGCCAAGTTTTGGAGAGCTTGAACTCACCATGAGACCACTTCATTGATTTGTGC	125
Oy	769	ATTCCTGTGTGATGTGACTATTGTGATAGCATCAAAGATATCGGGGGCAACAGCTCTG	828

Db 126 ATCTGTGCTGATGTACTTGTGATAGCATCAAAAGCTATCGGGGGGCAACAGCTCTG 185  
 QY 829 AACTTATTCGTACATGGGTCCATGAAAGCATCTTGGAGAATCTTAATTAAGACAGA 888  
 Db 186 AACTTATTCGTACATGGGTCCATGAAAGCATCTTGGAGAATCTTAATTAAGACAGA 245  
 QY 889 TATCAATTCCTGAGGAGCTGGCTTACCAAGAAAGCTGAGCTTGTTCAGAGACCTAAT 948  
 Db 246 TACCAAAATTCCTGAGGAGCTGGCTTACCAAGAAAGCTGAGCTTGTTCAGAGACCTAAT 305  
 QY 949 GTCAATTCGATATTCCTGAGGAGCTGGCTTACCAAGAAAGCTGAGCTTGTTCAGAGACCTAAT 1008  
 Db 306 GTCAATTCGATATTCCTGAGGAGCTGGCTTACCAAGAAAGCTGAGCTTGTTCAGAGACCTAAT 365  
 QY 1009 TTCCTGTAAAGATATGCTTTCACAGAGATCG 1044  
 Db 366 TTCCTGTAAAGATATGCTTTCACAGAGATCG 401

RESULT 15  
 AM562788/c 225 bp mRNA linear EST 10-MAR-2000  
 LOCUS 660065H06.X5 660 - Mixed stages of anther and pollen Zea mays cDNA,  
 DEFINITION mRNA sequence.

ACCESSION AM562788  
 VERSION AM562788.1 GI:7216666  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 225)  
 Walbot,V.  
 Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL Unpublished (1999)  
 COMMENT Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 723 8221  
 Email: walbot@stanford.edu  
 Plate: 660065 row: H column: 06.

FEATURES  
 source  
 1..225  
 /organism="Zea mays"  
 /cultivar="Oh1043"  
 /db\_xref="taxon:4577"  
 /clone\_lib="660 - Mixed stages of anther and pollen"  
 /tissue\_type="whole premeiotic anthers to pollen shed"  
 /dev\_stage="premeiotic anthers to pollen shed"  
 /lab\_host="X10LR"  
 /note="Organ: anthers; Vector: Lambda Zap; Site\_1: EcoRI;  
 Site\_2: XhoI; Anther and pollen cDNA library.  
 Directionally sequenced with 5' end at the EcoRI site.  
 Created by Amie Franklin."

BASE COUNT 40 a 53 c 45 g 87 t  
 ORIGIN  
 Query Match 12.8%; Score 187; DB 9; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-33;  
 Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 GAAGATGGGTGACAAAGGCCATGAGAAAGATCAATCTGCCAAGATTAATCTCGCAA 1095  
 Db 187 GAAGATGGGTGACAAAGGCCATGAGAAAGATCAATCTGCCAAGATTAATCTCGCAA 128  
 QY 1096 GGAAGACTCGAGTCCTTTTCAAGCCAACTGCGCCACATCAGCAGCGCTAAAGCGGAG 1155  
 Db 127 GGAAGACTCGAGTCCTTTTCAAGCCAACTGCGCCACATCAGCAGCGCTAAAGCGGAG 68

QY 1156 GAGACTTCGATTAACACAGCAGCTGCGACACAAGAAACAAAGCTGTGGAAG 1215  
 Db 67 GAGACTTCGATTAACACAGCAGCTGCGACACAAGAAACAAAGCTGTGGAAG 8  
 QY 1216 AAGAAAT 1222  
 Db 7 AAGAAAT 1

RESULT 16  
 AM288831 470 bp mRNA linear EST 16-JAN-2000  
 LOCUS 707010F11.y2 707 - Mixed adult tissues from Walbot lab (SK) Zea  
 DEFINITION mays cDNA, mRNA sequence.

ACCESSION AM288831  
 VERSION AM288831.1 GI:6695753  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 470)  
 Walbot,V.  
 Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL Unpublished (1999)  
 COMMENT Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 723 8221  
 Email: walbot@stanford.edu  
 Plate: 707010 row: F column: 11.

FEATURES  
 source  
 1..470  
 /organism="Zea mays"  
 /cultivar="W23"  
 /db\_xref="taxon:4577"  
 /clone\_lib="707 - Mixed adult tissues from Walbot lab (SK  
 )"  
 /tissue\_type="tassel, kernel, silk, husk, root, leaf"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: tassel, kernel, silk, husk, root, leaf;  
 Vector: pGAD10; Site\_1: EcoRI; cDNA library from fully  
 differentiated maize tissues from an active Mutator  
 plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,  
 husk, root, leaf). Unidirectionally cloned."

BASE COUNT 102 a 110 c 108 g 150 t  
 ORIGIN  
 Query Match 10.1%; Score 148; DB 9; Length 470;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-24;  
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 717 CAAGCTTTGAGAGAGCTGTAACCTCACCATGACACCTATCTATTTGTCATCTG 776  
 Db 14 CAAGCTTTGAGAGAGCTGTAACCTCACCATGACACCTATCTATTTGTCATCTG 73  
 QY 777 TGGATGTGATATGTGATAGCATCAAGGATATGGGGGGCAACACAGCTGAACTTAT 836  
 Db 74 TGGATGTGATATGTGATAGCATCAAGGATATGGGGGGCAACACAGCTGAACTTAT 133  
 QY 837 TCGTCAACATGGGTCCATPACAAAGCATC 864  
 Db 134 TCGTCAACATGGGTCCATPACAAAGCATC 161

RESULT 17  
 AM562518/c 363 bp mRNA linear EST 10-MAR-2000  
 LOCUS AM562518

DEFINITION 660065H06.x3 660 - Mixed stages of anther and pollen Zea mays cDNA,  
mRNA sequence.  
ACCESSION AWS62518  
VERSION AWS62518.1 GI:7216396  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 363)  
REFERENCE  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 660065 row: H column: 06.  
Location/Qualifiers  
1. 363  
/organism="Zea mays"  
/cultivar="Ohio43"  
/db\_xref="taxon:4577"  
/clone\_lib="660 - Mixed stages of anther and pollen"  
/tissue\_type="whole premeiotic anthers to pollen shed"  
/dev\_stage="premeiotic anthers to pollen shed"  
/lab\_host="XL0LR"  
/note="Organ: anthers; Vector: Lambda Zap; Site\_1: ECORI;  
Site\_2: XhoI; Anther and pollen cDNA library.  
Directionally sequenced with 5' end at the EcoRI site.  
Created by Ami Franklin."  
BASE COUNT 96 a 82 c 74 g 111 t  
ORIGIN

Query Match 8 9%; Score 130; DB 9; Length 363;  
Best Local Similarity 100.0%; Pred. No. 2.7e-20;  
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1141 CCGCTAAAGGAGAGACTTCGATTAACAAGCAAGCAGCTCGGACAGAAACA 1200  
|||||  
Db 301 CCGCTAAAGGAGAGACTTCGATTAACAAGCAAGCAGCTCGGACAGAAACA 242  
|||||

QY 1201 AAGCGTGTGGAAAGAATAATCTTGATGCTTGATGACACTACGACGAAG 1260  
|||||  
Db 241 AAGCGTGTGGAAAGAATAATCTTGATGCTTGATGACACTACGACGAAG 182  
|||||

QY 1261 CAGCGGTGGC 1270  
|||||  
Db 181 CAGCGGTGGC 172  
|||||

RESULT 18  
BF727781 232 bp mRNA linear EST 09-JAN-2001  
LOCUS BF727781  
DEFINITION 1000052F12.x2 1000 - Unigene 1 from Maize Genome Project Zea mays  
CDNA, mRNA sequence.  
ACCESSION BF727781  
VERSION BF727781.1 GI:12045642  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 232)  
REFERENCE  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY

JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 1000052 row: F column: 12.  
Location/Qualifiers  
1. 232  
/organism="Zea mays"  
/db\_xref="dbEST:707010C02.y1"  
/db\_xref="taxon:4577"  
/clone\_lib="1000 - Unigene 1 from Maize Genome Project"  
/note="This library represents the unique ESTs found in  
the first round of EST sequencing at Stanford University  
for the maize genome project. Sequences are present from  
libraries 486, 487, 496, 603, 605, 614, 618, 660, 683  
, 687, 707, and 945. Contigs were assembled using TIGR's  
CAP program and a representative EST from each contig was  
selected for the Unigene set. All singlets were also  
selected."  
BASE COUNT 64 a 42 c 57 g 69 t  
ORIGIN

Query Match 7.0%; Score 103; DB 10; Length 232;  
Best Local Similarity 100.0%; Pred. No. 3.9e-14;  
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 755 TCATTGATTTTGTCATCTGTCGATGTCACATTTTGATGATGATCAAGGTATCGGG 814  
|||||  
Db 55 TCATTGATTTTGTCATCTGTCGATGTCACATTTTGATGATGATCAAGGTATCGGG 114  
|||||

QY 815 GGCAACAGCTCTGAACCTATTGCTCACATGGGTCACATGA 857  
|||||  
Db 115 GGCAACAGCTCTGAACCTATTGCTCACATGGGTCACATGA 157  
|||||

RESULT 19  
AT834484/c 553 bp mRNA linear EST 02-FEB-2000  
LOCUS AT834484/c  
DEFINITION 606068G09.x1 606 - Ear tissue cDNA library from Schmidt lab Zea  
mays cDNA, mRNA sequence.  
ACCESSION AT834484  
VERSION AT834484.1 GI:5468693  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 553)  
REFERENCE  
AUTHORS Walbot,V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 606068 row: G column: 09.  
Location/Qualifiers  
1. 553  
/organism="Zea mays"  
/cultivar="Ohio43"  
/db\_xref="taxon:4577"  
/clone\_lib="606 - Ear tissue cDNA library from Schmidt  
lab"  
/tissue\_type="mixed"



```
/dev_stage="ear length from 0.5 cm - 2.0 cm"
/lab_host="XLOLR (Stratagene)"
/notes="Organ: Immature ear; Vector: pBK-CMV; Site_1: EcoRI
; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
lab"
BASE COUNT      140 a      112 c      106 g      195 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 68; DB 9; Length 553;
Pred. No. 3.9e-12;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1036 GAAGATCGGTCGCAAGGCCATGAGAGATCAATCTCCAGAGTAATGTCGCCAA 1095
|||||
Db 512 GAAGATCGGTCGCAAGGCCATGAGAGATCAATCTCCAGAGTAATGTCGCCAA 453
|||||
OY 1096 GGAAGACTGAGTCTTTTTCAGAGCCACTGCC 1128
|||||
Db 452 GGAAGACTGAGTCTTTTTCAGAGCCACTGCC 420
|||||

RESULT 20
LOCUS      AM147048      126 bp      mRNA      linear      EST 03-NOV-1999
DEFINITION 707010C02.y1 707 - Mixed adult tissues from Walbot lab (SK) Zea
ACCESSION  AM147048
VERSION     AM147048.1  GI:6194944
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 126)
REFERENCE  1
AUTHORS   Walbot V.
TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
          University
JOURNAL    Unpublished (1999)
COMMENT    Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 707010 row: C column: 02.

FEATURES
source
1..126
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
BASE COUNT      29 a      31 c      26 g      40 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 68; DB 9; Length 126;
Pred. No. 3.7e-06;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 797 GCATCAAGGTATCGGGGCAACAGCTGTGAACCTATTGTCACATGCGTCATAG 856
|||||
Db 1 GCATCAAGGTATCGGGGCAACAGCTGTGAACCTATTGTCACATGCGTCATAG 60
|||||
```

```
OY 857 AAGCATC 864
|||||
Db 61 AAGCATC 68
|||||

RESULT 21
LOCUS      AM147048/c      126 bp      mRNA      linear      EST 03-NOV-1999
DEFINITION 707010C02.y1 707 - Mixed adult tissues from Walbot lab (SK) Zea
ACCESSION  AM147048
VERSION     AM147048.1  GI:6194944
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 126)
REFERENCE  1
AUTHORS   Walbot V.
TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
          University
JOURNAL    Unpublished (1999)
COMMENT    Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 707010 row: C column: 02.

FEATURES
source
1..126
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
BASE COUNT      29 a      31 c      26 g      40 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 65; DB 9; Length 126;
Pred. No. 1.8e-05;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1172 CAAGCAAGCGACCTGCGACACAGAAACAAAGCGTGTGCAAGAGAAATATCTTGA 1231
|||||
Db 126 CAAGCAAGCGACCTGCGACACAGAAACAAAGCGTGTGCAAGAGAAATATCTTGA 67
|||||
OY 1232 TGGTT 1236
|||||
Db 66 TGGTT 62
|||||

RESULT 22
LOCUS      A1947478      225 bp      mRNA      linear      EST 19-AUG-1999
DEFINITION 614047B01.x1 614 - root cDNA library from Walbot lab Zea mays cDNA,
ACCESSION  A1947478
VERSION     A1947478.1  GI:5739683
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 225)  
Walbot, V.  
Maize ESTs from various cDNA libraries sequenced at Stanford  
University  
Unpublished (1999)  
Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 723 8221  
Email: walbot@stanford.edu  
Plate: 614047 row: B column: 01.  
Location/Qualifiers

FEATURES  
Source

1..225  
/organism="Zea mays"  
/cultivar="W23"  
/db\_xref="taxon:4577"  
/clone\_id="614 - root cDNA library from Walbot Lab"  
/tissue\_type="root"  
/dev\_stage="3-4 days old"  
/lab\_host="XLDLR"  
/note="Organ: root; Vector: pBluescriptII SK+; Site\_1:  
ECORI; Site\_2: XhoI; 3-4 days old root tissue from Walbot  
Lab (LM)"  
BASE COUNT 66 a 54 c 40 g 65 t  
ORIGIN

Query Match 3.9%; Score 57; DB 9; Length 225;  
Best Local Similarity 100.0%; Pred. No. 0.00076;  
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1185 TGGCAACAAGAAACAAAGCTGTGCAAGAAATATCTTGATGATGATCT 1241  
|||||  
Db 225 TGGCAACAAGAAACAAAGCTGTGCAAGAAATATCTTGATGATGATGATCT 169

RESULT 23  
LOCUS BG406288 346 bp mRNA linear EST 28-NOV-2001  
DEFINITION sac29h02.y1 Gm-cl051 glycine max cDNA clone GENOME SYSTEMS CLONE  
ID: Gm-cl051-3819 5' similar to TR:Q9SWS9 Q9SWS9 RIBOSOMAL PROTEIN  
S26.; mRNA sequence.  
ACCESSION BG406288  
VERSION BG406288.1 GI:13312637  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 346)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna  
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
R., Riter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
R., Waterston, R. and Wilson, R.  
Public soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
Putative full length read  
vector to vector length is 347 This clone is available through:  
ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL

FEATURES  
Source

35801 For further information call: (800)-533-4363 or contact via  
email: ccu@resgen.com.  
Location/Qualifiers

1..346  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone\_id="GENOME SYSTEMS CLONE ID: Gm-cl051-3819"  
/clone\_id="Gm-cl051"  
/tissue\_type="floral meristematic mRNA"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:  
XhoI; The cDNA library was constructed from floral  
meristematic mRNA provided by Dr. Halina Knap of Clemson  
University. Complementary DNA was synthesized from mRNA  
using a primer consisting of a poly(dT) sequence with a  
XhoI restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."  
BASE COUNT 107 a 49 c 73 g 117 t  
ORIGIN

Query Match 2.8%; Score 41; DB 10; Length 346;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1423 TTTCAGTAAATAAAAA 1463  
|||||  
Db 304 TTTCAGTAAATAAAAA 344

RESULT 24  
LOCUS AL121294/c 468 bp mRNA linear EST 25-FEB-2000  
DEFINITION AL121294  
DKFZp762L016.t1 762 (synonym: hme12) Homo sapiens cDNA clone  
DKFZp762L016 5', mRNA sequence.  
ACCESSION AL121294  
VERSION AL121294.1 GI:5927295  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 468)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Wiemann, S.  
Bloecher, H., Boecher, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and  
EST (Bloecher, et al.)  
Unpublished (1999)  
Contact: Bloecher H  
MPS  
Am Klopferstritz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;  
sequenced by GMP (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No sl sequence available.  
This clone (DKFZp762L016) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers

FEATURES  
Source

1..468  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_id="DKFZp762L016"  
/clone\_id="762 (synonym: hme12)"  
/tissue\_type="melanoma (Mewo cell line)"  
/dev\_stage="adult"  
/lab\_host="DH10B"

BASE COUNT 100 a 110 c 104 g 154 t  
ORIGIN  
Query Match 2.8%; Score 41; DB 9; Length 466;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1423 TTGGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
DB 68 TTGGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 28  
|||||  
RESULT 25  
BI761155 642 bp mRNA linear EST 25-SEP-2001  
LOCUS 603043646F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5184154 5',  
DEFINITION mRNA sequence.  
ACCESSION BI761155  
VERSION BI761155.1 GI:157523733  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS 1 (bases 1 to 642)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM11459 row: k column: 12  
High quality sequence stop: 383.  
Location/Qualifiers  
1. 642  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5184131"  
/clone\_lib="NIH\_MGC\_116"  
/lab\_host="DH10B"  
/note="Organ: pooled colon, kidney, stomach; Vector:  
pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
source anonymous pool of 3 colons, age 26 yo male, 49 yo  
female, 71 yo male colon; 46 yo male kidney, and pool of 2  
stomachs, 62 yo male and 70 yo female. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.4 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
023. Note: this is a NIH MGC Library."  
BASE COUNT 234 a 148 c 151 g 109 t  
ORIGIN  
Query Match 2.8%; Score 41; DB 10; Length 642;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1423 TTGGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
DB 269 TTGGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 309  
|||||  
RESULT 26  
BI760815 822 bp mRNA linear EST 25-SEP-2001  
LOCUS 603043646F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5184154 5',  
DEFINITION mRNA sequence.  
ACCESSION BI760815  
VERSION BI760815.1 GI:157523933  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS 1 (bases 1 to 822)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM11459 row: 1 column: 11  
High quality sequence start: 3  
High quality sequence stop: 370.  
Location/Qualifiers  
1. 822  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5184154"  
/clone\_lib="NIH\_MGC\_116"  
/lab\_host="DH10B"  
/note="Organ: pooled colon, kidney, stomach; Vector:  
pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
source anonymous pool of 3 colons, age 26 yo male, 49 yo  
female, 71 yo male colon; 46 yo male kidney, and pool of 2  
stomachs, 62 yo male and 70 yo female. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.4 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
023. Note: this is a NIH MGC Library."  
BASE COUNT 301 a 205 c 211 g 105 t  
ORIGIN  
Query Match 2.8%; Score 41; DB 10; Length 822;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1423 TTGGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
DB 268 TTGGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 308  
|||||  
RESULT 27  
AM307211 202 bp mRNA linear EST 02-DEC-2001  
LOCUS sfs4b11.y1 Gm-c1009 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
DEFINITION Gm-c1009-3766 5', mRNA sequence.  
ACCESSION AM307211  
VERSION AM307211.1 GI:6719564  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 202)  
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna  
A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

603043646F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5184154 5',  
mRNA sequence.  
ACCESSION BI760815  
VERSION BI760815.1 GI:157523933  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS 1 (bases 1 to 822)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM11459 row: 1 column: 11  
High quality sequence start: 3  
High quality sequence stop: 370.  
Location/Qualifiers  
1. 822  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5184154"  
/clone\_lib="NIH\_MGC\_116"  
/lab\_host="DH10B"  
/note="Organ: pooled colon, kidney, stomach; Vector:  
pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
source anonymous pool of 3 colons, age 26 yo male, 49 yo  
female, 71 yo male colon; 46 yo male kidney, and pool of 2  
stomachs, 62 yo male and 70 yo female. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.4 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
023. Note: this is a NIH MGC Library."  
BASE COUNT 301 a 205 c 211 g 105 t  
ORIGIN  
Query Match 2.8%; Score 41; DB 10; Length 822;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1423 TTGGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
|||||  
DB 268 TTGGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 308  
|||||  
RESULT 27  
AM307211 202 bp mRNA linear EST 02-DEC-2001  
LOCUS sfs4b11.y1 Gm-c1009 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
DEFINITION Gm-c1009-3766 5', mRNA sequence.  
ACCESSION AM307211  
VERSION AM307211.1 GI:6719564  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 202)  
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna  
A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

TITLE  
JOURNAL  
COMMENT  
Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shiu, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: cou@resgen.com  
Seq primer: -40RP from Gibco.

# FEATURES

## source

1. 202  
Location/Qualifiers

/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl009-3766"  
/clone\_lib="Gm-cl009"  
/lab\_host="XLI0-Gold"  
/note="Vector: pBluescript II XR, Site\_1: EcoRI; Site\_2: XhoI; The mRNA was isolated from entire roots of 2-month-old 'Williams' plants that were greenhouse grown in 5-gallon pots. To suppress nodulation, Black Gold All-Purpose potting soil was supplemented with: 0.36g/L available phosphoric acid (P205), 20mg/L urea N, 0.16g/L S, 0.49mg/L B, 2.5mg/L Cu, 0.15g/L Fe, 13.53mg/L Mn, 0.26mg/L Mo, 14mg/L Zn, 20mg/L Ca, and the following nutrients in a slow-release form (Osmocote): 0.165g/L ammonia N, 0.185g/L nitrate N, 0.35g/L available phosphoric acid, and 0.35g/L soluble potash. No nodules were visible on the roots at harvest. Stratagene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. First strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA is hemimethylated. Stratagene's first-strand synthesis primer was used [GAGAGAGAGAGAGAGAGACTACTGCGAG(T)-18]. After second-strand synthesis, the cDNA ends were 'polished' with clone Pfu DNA polymerase, ligated to EcoRI adapters, and phosphorylated. The XhoI site within the first-strand synthesis primer was restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 400bp cutoff, using a SizeSeph 400 Spin column from Pharmacia. The column eluent was then ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) that had been digested with EcoRI and XhoI, and phosphorylated. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts. This library was constructed by Dr. Paul Keim and Dr. Virginia Corryell."

# BASE COUNT

102 a 22 c 22 g 56 t

Query Match 2.7%; Score 40; DB 9; Length 202;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 TTGAAGTAAAAA  
Db 127 TTGAAGTAAAAA  
163  
166

RESULT 28  
LOCUS AI373782 203 bp mRNA linear EST 16-FEB-1999  
DEFINITION q254g06.x1 NCI\_CGAP\_Kid11 Homo sapiens cDNA clone IMAGE:2030746 3', mRNA sequence.  
ACCESSION AI373782  
VERSION AI373782.1 GI:4153648

KEYWORDS  
SOURCE EST.  
ORGANISM human.  
Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 203)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

# JOURNAL

Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Greg Lennon, Ph.D.  
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LUNL at: www.bio.lnl.gov/bbrp/image/image.html  
Insert length: 304 Std Error: 0.00  
Seq primer: -40UP from Gibco.

# FEATURES

1. 203  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2030746"  
/clone\_lib="NCI\_CGAP\_Kid11"  
/lab\_host="DH10B"  
/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

# BASE COUNT

59 a 30 c 28 g 86 t

Query Match 2.7%; Score 40; DB 9; Length 203;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 TTGAAGTAAAAA  
Db 42 TTGAAGTAAAAA  
1463  
3

# RESULT 29

LOCUS AA678858/c 294 bp mRNA linear EST 02-DEC-1997  
DEFINITION ah05a03.s1 Gessler Wilms tumor Homo sapiens cDNA clone IMAGE:1155724 3', mRNA sequence.

ACCESSION AA678858  
VERSION AA678858.1 GI:2659380  
KEYWORDS EST.  
SOURCE human.  
ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 294)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maita, N., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.

# TITLE

WashU-NCI human EST Project  
Unpublished (1997)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108



BASE COUNT	ORIGIN
116 a	90 c 78 g 127 t
Query Match	2.7%; Score 40; DB 10; Length 411;
Best Local Similarity	100.0%; Pred. No. 3.3;
Matches	40: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1424 TTGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB	43 TTGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4
RESULT 32	
AA901800/c	LOCUS 522 bp mRNA linear EST 12-APR-1998
DEFINITION	NCC1A37 Confidial Neurospora crassa cDNA clone NCA3 3' end, mRNA
ACCESSION	sequence.
VERSION	AA901800
KEYWORDS	AA901800.1 GI:3045777
SOURCE	EST.
ORGANISM	Neurospora crassa.
REFERENCE	Neurospora crassa
AUTHORS	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariales; Sordariaceae; Neurospora. 1 (bases 1 to 522) Nelson,M.A., Kang,S., Braun,E.L., Crawford,M.E., Dolan,P.L., Leonard,P.M., Mitchell,J., Armijo,A.M., Bean,L., Blueses,E., Cushing,T., Erclet,A., Fleharty,M., Gorman,M., Judson,K., Miller,R., Ortega,J., Pavlova,I., Perez,J., Todisco,S., Trujillo,R., Valentine,J., Wells,A., Werner-Washburne,M., Yazzie,S. and Natvig D.O. Expressed sequences from confidial, mycelial, and sexual stages of Neurospora crassa Fungal Genet. Biol. 21, 348-363 (1997)
TITLE	Contact: Natvig,D.O./Nelson,M.A.
JOURNAL	Department of Biology
MEDLINE	University of New Mexico
COMMENT	Casteller Hall, Albuquerque, NM 87131, USA Tel: 505 277 3411 Fax: 505 277 0304 Email: npebiology.unm.edu Putative VACUOLAR ATP SYNTHASE CATALYTIC SUBUNIT A, based upon BLASTY search. Deposited in GSDb at the National Center for Genome Resources with accession GSDb:S:1147857

```

FEATURES
Source
Seq primer: 17.
Location/Qualifiers
1..522
/organism="Neurospora crassa"
/strain="74-OR23-IV A"
/db_xref="taxon:5141"
/clone="NCIA3"
/clone_id="Conidia"
/sex="Mating type A"
/tissue_type="Conidia"
/dev_stage="germinating conidia"
/note="mRNA isolated from germinating conidia, grown in 1x
Vogel's, 28 sucrose for 4.5 hours. cDNA directionally
cloned into pBluescript SK(-) using the Uni-ZAP XR vector
system (Stratagene, La Jolla, CA)."

BASE COUNT
110 a 143 c 65 g 204 t

Query Match
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1424 TTGACGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 87 TTGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 48

RESULT 33
LOCUS BF134331 741 bp mRNA linear EST 24-OCT-2000
DEFINITION 60178431471 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4012178 5',
mRNA sequence.
ACCESSION BF134331
VERSION BF134331.1 GI:10973371
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 741)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strusberg, Ph.D.
Email: cga@b5-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LMNL)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.lnl.gov
Plate: L1AM9252 row: 1 column: 03
High quality sequence stop: 231.
Location/Qualifiers
1..741
/organism="Mus musculus"
/strain="CEBCH II"
/db_xref="taxon:10090"
/clone="IMAGE:4012178"
/clone_id="NCI_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/note="Organ: Lung; Vector: pCMV-SPORT6; Site:1: NotI;
Site-2: SalI; transgenic model WNT-1, expression driven by
MTVY-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT
251 a 181 c 170 g 139 t

Query Match
Best Local Similarity 100.0%; Pred No. 2.3;
2.7%; Score 40; DB 10; Length 741;
100.0%;

```

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1424 TCGAAGTAAAAA 1463  
 Db 192 TTGAAGTAAAAA 231

## RESULT 34

AM568974

LOCUS 109 bp mRNA linear EST 03-DEC-2001  
 DEFINITION s173909.y1 Gm-cl031 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-cl031-377 5', mRNA sequence.

## ACCESSION

AM568974

## VERSION

AM568974.1

## KEYWORDS

EST

## SOURCE

soybean.

## ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

## REFERENCE

1 (bases 1 to 109)

## AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpel, J., Corryell, V., Khanna,  
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,  
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk,  
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann,  
 R., Waterston, R. and Wilson, R.

## TITLE

Public Soybean EST Project

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: cdu@resgen.com  
 High quality sequence stop: 67.

## FEATURES

source

1. 109  
 Location/Qualifiers  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl031-377"  
 /clone\_lib="Gm-cl031"  
 /tissue\_type="Williams seedlings, minus the cotyledons"  
 /lab\_host="D10B"  
 /note="Vector: pBluescript II XR; Site\_1: EcoRI; Site\_2:  
 XhoI; This cDNA library was constructed from mRNA isolated  
 from whole 'Williams' seedlings, minus the cotyledons,  
 which were propagated on paper towels with distilled water  
 for 5 days, incubated at 40 degrees C for 1 hour. The  
 cotyledons were removed and the remaining tissue was  
 flash-frozen in liquid nitrogen. StrataGene's cDNA  
 Synthesis Kit (catalog number 200401) was used to  
 synthesize the cDNA. First-strand synthesis was performed  
 with 5-methyl dCTP, hence the ligated cDNA was  
 hemimethylated. A modification of StrataGene's  
 first-strand synthesis primer was used. An 'anchor'  
 nucleotide (V=A, C, or G) was added to the 3' end of the  
 primer (GAGACAGACAGACAGACAGACTGTCGAC(T)18V) to anchor  
 the primer at the 5' end of the poly(A) tract. After  
 second strand synthesis, the cDNA ends were filled in with  
 cloned Pfu DNA, ligated to EcoRI adapters and  
 subsequently phosphorylated. The cDNA was then  
 precipitated and redissolved in sterile, RNase-  
 dnase-free water. The XhoI site within the first-strand  
 synthesis primer was then restricted by digestion with  
 XhoI from Promega (400/ul); all XhoI sites in the cDNA  
 would be protected by their hemimethylated status. The  
 cDNA constructs were size-fractionated with a 500 bp

BASE COUNT 67 a 7 c 14 g 21 t

## ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 12;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1425 TGAAGTAAAAA 1463  
 Db 52 TGAAGTAAAAA 90

## RESULT 35

BI319195

LOCUS 150 bp mRNA linear EST 23-JUL-2001  
 DEFINITION 949026H11.y1 949 - Juvenile leaf and shoot cDNA from Steve Moose

## ACCESSION

BI319195

## VERSION

BI319195.1

## KEYWORDS

EST

## SOURCE

zea mays.

## ORGANISM

zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 150)

## REFERENCE

Walbot, V.

## AUTHORS

Maize ESTs from various cDNA libraries sequenced at Stanford

## TITLE

University

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 949026 row: H column: 11.

## FEATURES

source

1. 150  
 Location/Qualifiers  
 /organism="Zea mays"  
 /cultivar="W64A"  
 /db\_xref="taxon:4577"  
 /clone\_lib="949 - Juvenile leaf and shoot cDNA from Steve  
 Moose"  
 /tissue\_type="Immature leaf primordium and vegetative  
 meristem"  
 /dev\_stage="4 stages from 3-13 days after imbibing"  
 /lab\_host="E. coli XL10R"  
 /note="Organ: Juvenile vegetative shoots; Vector:  
 pMD-GAL4-2.1; Site\_1: EcoRI; Site\_2: XhoI; Equal amounts  
 of total RNA by weight from 4 tissue sources (see below)  
 were pooled, polyA+ RNA isolated, and cDNA synthesized for  
 EcoRI (5') and XhoI (3') directional cloning into lambda  
 HybriZap vector from StrataGene. Tissue Sources: 1. Whole  
 shoots 3 days after sowing/imbibing in wet soil. 2. Basal  
 portions of developing leaves 1-5, primordia from 6-8, and  
 the vegetative apex. 3. Non-green portions of developing  
 leaves 4-5 and the vegetative apex, including adult leaf  
 primordia. 9 days after sowing. 4. Partially expanded and  
 greening leaves 4-5 at 13 days after sowing."

## BASE COUNT

86 a 17 c 14 g 33 t

## ORIGIN

Query Match 2.7%; Score 39; DB 10; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 |||||||||||||||||||||||||||||||||||||||||  
 Db 95 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 133

RESULT 36  
 BF749502 187 bp mRNA linear EST 10-JAN-2001  
 LOCUS IL0-BN0422-121000-500-b06 BN0422 Homo sapiens CDNA, mRNA sequence.  
 DEFINITION BF749502  
 ACCESSION BF749502.1 GI:12076178  
 VERSION EST.  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 187)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/ICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL0&t=IL0-BN0422-  
 121000-500-b06&t3=2000-10-12&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 20  
 High quality sequence stop: 186.  
 Location/Qualifiers  
 1. 187  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="BN0422"  
 /dev\_stage="Adult"  
 /note="Organ: breast, normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 78 a 25 c 36 g 48 t  
 ORIGIN

Query Match 2.7%; Score 39; DB 10; Length 187;  
 Best Local Similarity 100.0%; Pred. No. 9;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 |||||||||||||||||||||||||||||||||||||||||  
 Db 142 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180

## RESULT 37

BF749503 201 bp mRNA linear EST 10-JAN-2001  
 LOCUS IL0-BN0422-121000-500-b09 BN0422 Homo sapiens CDNA, mRNA sequence.  
 DEFINITION BF749503  
 ACCESSION BF749503.1 GI:12076179  
 VERSION EST.  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 201)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/ICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL0&t=IL0-BN0422-  
 121000-500-b09&t3=2000-10-12&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 34  
 High quality sequence stop: 200.  
 Location/Qualifiers  
 1. 201  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="BN0422"  
 /dev\_stage="Adult"  
 /note="Organ: breast, normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

JOURNAL MEDLINE  
 COMMENT

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

## FEATURES

## source

1. 201  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="BN0422"  
 /dev\_stage="Adult"  
 /note="Organ: breast, normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 82 a 27 c 41 g 51 t  
 ORIGIN

Query Match 2.7%; Score 39; DB 10; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 8.6;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 |||||||||||||||||||||||||||||||||||||||||  
 Db 158 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 196

## RESULT 38

BF764516 226 bp mRNA linear EST 12-JAN-2001  
 LOCUS R01-CS0068-011100-011-g03 CS0068 Homo sapiens CDNA, mRNA sequence.  
 DEFINITION BF764516  
 ACCESSION BF764516.1 GI:12112416  
 VERSION EST.  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.



REFERENCE 1 (bases 1 to 226)  
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J., and Simpson, A. J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the PAPERP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=KCl&cl2=KCl-CS0068-011100-011-g03&cl3=2000-11-01&cl4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 25  
 High quality sequence stop: 186.

FEATURES  
 source location/Qualifiers  
 1..226  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="CS0068"  
 /dev\_stage="Adult"  
 /note="Organ: colon\_est; Vector: puc18; site:1: SmaI; site:2: SmaI: A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 122 a 27 c 33 g 44 t

ORIGIN

Query Match 2.7%; Score 39; DB 10; Length 226;  
 Best Local Similarity 100.0%; Pred. No. 8;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1425 TGAAGTAAAAA  
 Db 132 TGAAGTAAAAA

RESULT 39  
 AM234203 320 bp mRNA linear EST 02-DEC-2001  
 LOCUS sf22d11.y1 Gm-c1028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-c1028-718 5', mRNA sequence.  
 ACCESSION AM234203  
 VERSION AM234203.1 GI:6566560  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 1 (bases 1 to 320)  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, R., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project

JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Parkway  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: coute@resgen.com  
 Insert Length: 394 Std Error: 0.00  
 High quality sequence stop: 306.

FEATURES  
 source location/Qualifiers  
 1..320  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1028-718"  
 /clone\_lib="Gm-c1028"  
 /tissue\_type="roots of 'Supernod' plants"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II XR; Site:1: EcoRI; site:2: XhoI; The mRNA was isolated from roots of Glycine max 'Supernod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicus, strain USDA110 prior to harvest. StrataGene's cDNA synthesis kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of StrataGene's first-strand synthesis primer was used. An 'anchor' nucleotide (V-A-C, or G) was added to the 3' end of the primer (GAGAGAGAGAGAGAGAGACATGTCGAG(T)18V) to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GIBCOBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into StrataGene's pBluescript II XR predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by StrataGene. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 110 a 38 c 66 g 106 t

ORIGIN

Query Match 2.7%; Score 39; DB 9; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 6.5;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1425 TGAAGTAAAAA  
 Db 275 TGAAGTAAAAA

RESULT 40  
 A2953545 365 bp DNA linear GSS 27-APR-2001  
 LOCUS A2953545/c  
 DEFINITION 2M0218A19R Mouse 10kb plasmid UUGCM library Mus musculus genomic  
 ACCESSION A2953545  
 VERSION A2953545.1 GI:13824772  
 KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 365)  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.  
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0218 row: A column: 19  
 Seq primer: CACACAGAAACACGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 365.  
 Location/Qualifiers  
 1..365  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUCG2M0218A19"  
 /clone\_lib="Mouse 10kb plasmid UUCG2M library"  
 /sex="Female"  
 /lab\_host="E. coli strain XL10-gold, T1-resistant, F-"  
 /note="Vector: PMD29v; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (g9147321419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 85 a 65 c 40 g 175 t  
 ORIGIN

Query Match 2.7%; Score 39; DB 12; Length 365;  
 Best Local Similarity 100.0%; Pred. No. 6;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1425 TGAAGTAAAAA  
 Db 150 TGAAGTAAAAA  
 RESULT 41  
 AA121171/c  
 LOCUS AA121171 431 bp mRNA linear EST 19-MAY-1997  
 DEFINITION 2188605.s1 Stratagene colon (#937204) Homo sapiens cDNA clone IMAGE:511736 3' similar to gb:W65066 CAMP-DEPENDENT PROTEIN KINASE TYPE I-ALPHA REGULATOR (HUMAN);, mRNA sequence.  
 ACCESSION AA121171  
 VERSION AA121171.1 GI:1678714  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 431)  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 AUTHORS Hillier,L., Lennon,G., Becker,M., Donald,M.F., Chippelli,B., Chissee,S., Dietrich,N., Dubuque,T., Favellio,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevasakis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M. Generation and analysis of 280,000 human expressed sequence tags  
 JOURNAL Genome Res. 6 (9), 807-828 (1996)  
 MEDLINE 97044478  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 1257 Std Error: 0.00  
 Seq primer: -40M13 fwd. from Amersham  
 High quality sequence stop: 93.  
 Location/Qualifiers  
 1..431  
 /organism="Homo sapiens"  
 /db\_xref="GDB:3844231"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:511736"  
 /clone\_lib="Stratagene colon (#937204)"  
 /tissue\_type="tumor"  
 /cell\_line="T84 carcinoma cell line"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /note="Organ: colon; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT; T-84 colonic epithelial cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACACAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

BASE COUNT 84 a 103 c 125 g 107 t 12 others  
 ORIGIN

Query Match 2.7%; Score 39; DB 9; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 5.4;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1425 TGAAGTAAAAA  
 Db 46 TGAAGTAAAAA  
 RESULT 42  
 AA121150/c  
 LOCUS AA121150 432 bp mRNA linear EST 19-MAY-1997  
 DEFINITION 2188607.s1 Stratagene colon (#937204) Homo sapiens cDNA clone IMAGE:511693 3' similar to gb:D00761 PROTEASOME COMPONENT C5 (HUMAN);, mRNA sequence.  
 ACCESSION AA121150  
 VERSION AA121150.1 GI:1678693  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 432)  
 AUTHORS Hillier,L., Lennon,G., Becker,M., Donald,M.F., Chippelli,B., Chissee,S., Dietrich,N., Dubuque,T., Favellio,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevasakis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M. Generation and analysis of 280,000 human expressed sequence tags  
 JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 97044478  
COMMENT

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 1006 Std Error: 0.00  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 354.

## FEATURES

source

```
1..432
/organism="Homo sapiens"
/db_xref="GDB:3844188"
/db_xref="taxon:9606"
/clone="IMAGE:511693"
/clone_lib="Stratagene colon (#937204)"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: colon; Vector: pBluescript SK-; Site: 1;
EcoRI; Site: 2; XhoI; Cloned unidirectionally. Primer:
Oligo dt. T-84 colonic epithelial cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor
sequence: 5' GAATCGGCGACGAG-3' -3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3'."
BASE COUNT 105 a 111 c 79 g 137 t
ORIGIN
```

## Query Match

Best Local Similarity 100.0%; Pred. No. 5.4;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
Db 39 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1

## RESULT 43

BM357517 444 bp mRNA linear EST 30-JAN-2002  
LOCUS BM357517/c

DEFINITION Triphysaria versicolor root-tip, early DMO-induced

transcript cDNA library Triphysaria versicolor cDNA, mRNA sequence.

ACCESSION BM357517

VERSION BM357517.1 GI:18426469

KEYWORDS EST.

SOURCE Triphysaria versicolor.

ORGANISM Triphysaria versicolor.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;

Asteridae; euasterids I; Lamiales; Orobanchaceae; Triphysaria.

1 (bases 1 to 444)

REFERENCE Matvienko, M., Torres, M.J. and Yoder, J.I.

Transcriptional responses to the hemiparasitic plant Triphysaria

versicolor to host plant signals

Journal Plant Physiol. 127 (1), 272-282 (2001)

JOURNAL MEDLINE 21437952

COMMENT Contact: Yoder, J.I.

John I. Yoder Research Lab, Dept. of Vegetable Crops

University of California at Davis

137 Assumndson Hall, One Shields Drive, Davis, CA 95616, USA

Tel: 530 752 1741

Fax: 530 752 9659

Email: jyoder@ucdavis.edu

High quality sequence stop: 445.

Location/Qualifiers

1..444

/organism="Triphysaria versicolor"

/db\_xref="taxon:96093"

/clone\_lib="Triphysaria versicolor root-tip, early

DMO-induced transcript cDNA library"

/tissue\_type="root-tips ~5mm length"

/dev\_stage="3-4 weeks growth"

/lab\_host="E. coli"

/note="Vector: pCR2.1 TA Cloning System, Invitrogen,

Carlsbad, CA; PCR-based suppression subtractive

hybridization cDNA library"

BASE COUNT 120 a 119 c 60 g 145 t

ORIGIN

Query Match 2.7%; Score 39; DB 10; Length 444;

Best Local Similarity 100.0%; Pred. No. 5.3;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463

Db 66 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 28

## RESULT 44

AI362270/c

LOCUS AI362270

DEFINITION

qy50e11.x1 NCI-CGAP\_Brn23 Homo sapiens cDNA clone IMAGE:2015468 3'

similar to gb:M54927 MYELIN PROTEOLIPID PROTEIN (HUMAN); mRNA

sequence.

ACCESSION AI362270

VERSION AI362270.1 GI:4113891

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 448)

REFERENCE NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/cgi/gap.

National Cancer Institute / National Institute of Neurological

Disorders and Stroke, Brain Tumor Genome Anatomy Project

(CGAP/BRGAP), Tumor Gene Index

Unpublished (1998)

Contact: Robert Strausberg, Ph.D.

Email: cgaaps-remail.nih.gov

Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,

Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/BLN at:

www.bio.lnl.gov/dbp/image/image.html

Insert length: 1090 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 285.

Location/Qualifiers

1..448

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2015468"

/clone\_lib="NCI CGAP\_Brn23"

/tissue\_type="glioblastoma (pooled)"

/lab\_host="DH10B"

/note="Organ: Brain; Vector: p773D-Pac (Pharmacia) with a

modified polylinker; Site: 1; Not I; Site: 2; Eco RI; 1st

strand cDNA was primed with a Not I - Oligo(dt) primer [5'

TGTTCACATCTGAAGTGGAGGCGGCGATATCTTTTGTTTTGTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified p773 vector.

Library is normalized, and was constructed by Bento

Soares and M.Fatima Bonaldo."

BASE COUNT 127 a 86 c 72 g 163 t

ORIGIN

Query Match 2.7%; Score 39; DB 9; Length 448;

Best Local Similarity 100.0%; Pred. No. 5.3;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 ||||||||||||||||||||||||||||||||||||||||  
 DB 42 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4

RESULT 45  
 AL501532/c 593 bp mRNA linear EST 04-JAN-2001  
 LOCUS AL501532 Hordeum vulgare Barke roots Hordeum vulgare cDNA clone  
 DEFINITION HM041040 3', mRNA sequence.  
 ACCESSION AL501532  
 VERSION AL501532.1 GI:12027747  
 KEYWORDS EST.  
 SOURCE barley.  
 ORGANISM Hordeum vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
 ; Triticeae; Hordeum.

REFERENCE  
 AUTHORS Michael, W., Weschke, W., Pleissner, K.-P. and Graner, A.  
 TITLE EST sequencing and analysis in barley  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Michael, W.  
 Institute for Plant Genetics and Crop Plant Research  
 Corrensstr.3, D-06466 Gatersleben, Germany  
 Email: michael@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de  
 Seq primer: M13un1 primer for 3'end.

FEATURES  
 source  
 Location/Qualifiers  
 1..593  
 /organism="Hordeum vulgare"  
 /cultivar="Barke"  
 /db\_xref="taxon:4513"  
 /clone\_1lb="HM041040"  
 /clone\_1lb="Hordeum vulgare Barke roots"  
 /tissue\_type="roots"  
 /lab\_host="XIOLR"  
 /note="Vector: Plasmid pBK-CMV; Site.1: EcoRI; Site.2:  
 XhoI; mRNA was made from roots of spring barley variety  
 'Barke', a high quality malting variety. Roots were grown  
 for two days on filter paper at room temperature. Cloning  
 sites: EcoRI (5'-end of cDNA) and XhoI (3'-end of cDNA).  
 NOTE: Due to a cloning artefact caused by the kit, in most  
 cases the EcoRI site is NOT present, as well as the EcoRI  
 adapter. Average insert size is 1 kb. Sequence trimming:  
 Vector sequences and sequence ends were trimmed from the  
 5'- and 3'-end until a 50 bp window contains less than two  
 ambiguities. The maximum length was set to 700 bp"

BASE COUNT 165 a 122 c 114 g 191 t 1 others  
 ORIGIN

Query Match 2.7%; Score 39; DB 9; Length 593;  
 Best Local Similarity 100.0%; Pred.No. 4.5;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463  
 ||||||||||||||||||||||||||||||||||||||||  
 DB 42 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4

Search completed: November 5, 2002, 16:21.48  
 Job time : 1287.78 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 13:37:21 ; Search time 63 seconds  
(without alignments)  
668.206 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 379  
Sequence: 1 MGIGLTKLADNAPKAMKE.....SDKTSKAANKKTRAGKKK 379

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A.Geneseq\_032802:\*

```
1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	379	100.0	379	21	AAV95307
2	379	100.0	379	21	AAV95309
3	321	84.7	379	21	AAV95308
4	278	73.4	379	21	AAV95310
5	15	4.0	343	22	AB96644
6	14	3.7	385	22	AB963960
7	13	3.4	377	20	AAW92505
8	13	3.4	378	20	AAW92508
9	13	3.4	380	20	AAW92504
10	12	3.2	325	19	AAW59950
11	12	3.2	326	19	AAW59953

12	12	3.2	332	19	AAW59951
13	12	3.2	340	18	AAW24216
14	12	3.2	340	19	AAW79970
15	12	3.2	340	19	AAW59940
16	9	2.4	16	20	AAW92511
17	9	2.4	382	20	AAW92506
18	8	2.1	145	12	AAV10723
19	8	2.1	338	22	ABG08130
20	8	2.1	373	20	AAV03778
21	8	2.1	373	8	AAV70501
22	8	2.1	399	22	ABW59358
23	8	2.1	399	22	ABW5740
24	8	2.1	515	21	AAV95714
25	8	2.1	687	21	AAV58997
26	8	2.1	770	21	AAV43478
27	7	1.8	14	22	AAW68822
28	7	1.8	25	20	AAW92515
29	7	1.8	49	22	AAW90755
30	7	1.8	61	21	AAW34731
31	7	1.8	62	22	AAU06110
32	7	1.8	72	22	AAU02011
33	7	1.8	73	21	AAW05446
34	7	1.8	73	21	AAW07581
35	7	1.8	73	21	AAW08148
36	7	1.8	73	21	AAW08766
37	7	1.8	73	21	AAW36961
38	7	1.8	73	21	AAW52931
39	7	1.8	73	21	AAV59186
40	7	1.8	73	21	AAV59187
41	7	1.8	73	21	AAV59188
42	7	1.8	75	22	AAW63518
43	7	1.8	86	22	AAW84718
44	7	1.8	101	22	ABG17892
45	7	1.8	106	21	AAW05445

#### ALIGNMENTS

RESULT 1	
ID	AAV95307
AAV95307	standard; Protein; 379 AA.
AC	AAV95307;
XX	
DT	12-SEP-2000 (first entry)
XX	
DE	Maize Rad2/FEN-1 protein.
XX	
KW	Maize; Rad2/FEN-1; transgenic plant; male sterile plant; endonuclease; exonuclease; DNA repair; gene targeting.
XX	
OS	Zea mays.
XX	
PN	WO200036109-A1.
XX	
PD	22-JUN-2000.
XX	
PF	16-NOV-1999; 99WO-US27147.
XX	
PR	15-DEC-1998; 98US-0112332.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Mahajan PB;
XX	
DR	WPI; 2000-452026/39.
XX	
NR	N-PSDB; AAA27923.
XX	
PT	Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA recombination and repair in transgenic plants, e.g. for gene targeting and the production of male sterile plants -
XX	

PS Claim 11, Page 71-72; 85bp; English.  
XX  
CC The present sequence is that of maize RAD2/FEN-1, as deduced from  
CC a cDNA clone (see AAA27923) isolated from maize line B73 immature  
CC ear tissue. Rad2/FEN-1 is a structure specific endonuclease which  
CC under certain conditions also acts as an exonuclease. Rad2/FEN-1  
CC can be expressed in transgenic plant cells using conventional  
CC methods. The protein is involved in the regulation of DNA repair  
CC and recombination in plant systems and therefore may be used for  
CC improving gene targeting during further recombinant DNA protocols  
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential  
CC in DNA replication and nucleotide excision and repair reactions.  
CC The exolytic activity is involved in double strand break repair and  
CC end joining. The protein is also useful in strand exchange  
CC reactions during homologous recombination. These functions may be  
CC useful in gene targeting and in the production of male sterile  
CC plants. The efficacy of gene targeting can be improved by the  
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
CC be produced by the down regulation of Rad2/FEN-1 expression.  
XX  
SQ Sequence 379 AA;  
  
Query Match 100.0%; Score 379; DB 21; Length 379;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MGIGLTKLADNAPKAMKEQKFESEYFGKRIADVASMISYQFLIVGRTGMETLTNEAGE 60  
DB 1 MGIGLTKLADNAPKAMKEQKFESEYFGKRIADVASMISYQFLIVGRTGMETLTNEAGE 60  
QY 61 VTSHLQGMFNRTRIRLEAGIKPYVYFDGKPPDMKKQELAKRYSKRDATKDLTEAVEVGD 120  
DB 61 VTSHLQGMFNRTRIRLEAGIKPYVYFDGKPPDMKKQELAKRYSKRDATKDLTEAVEVGD 120  
QY 121 KDAIEKLSKRTVAVTROHNEDECKRLRLMGVPVVEAPSEAEACALCINDKVFPAVASED 180  
DB 121 KDAIEKLSKRTVAVTROHNEDECKRLRLMGVPVVEAPSEAEACALCINDKVFPAVASED 180  
QY 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAKVLELELTMDFIDLCILCGDYCDYSIK 240  
DB 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAKVLELELTMDFIDLCILCGDYCDYSIK 240  
QY 241 GIGGOTALKLIRHGSIESILENLNDRQIPEDMPIQEARRLFKEPNVTLDIPELKMTA 300  
DB 241 GIGGOTALKLIRHGSIESILENLNDRQIPEDMPIQEARRLFKEPNVTLDIPELKMTA 300  
QY 301 PDEGLISFLVKDNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
DB 301 PDEGLISFLVKDNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
QY 361 DKTSKAANKKTKAGGKK 379  
DB 361 DKTSKAANKKTKAGGKK 379  
  
RESULT 2  
AA95309  
ID AA95309 standard; Protein; 379 AA.  
XX  
AC AA95309;  
XX  
DT 12-SEP-2000 (first entry)  
XX  
DE Maize Rad2/FEN-1 protein.  
XX  
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;  
XX endonuclease; exonuclease; DNA repair; gene targeting.  
OS Zea mays.  
XX  
XX WO200036109-A1.  
PN  
XX  
PD 22-JUN-2000.

XX  
PF 16-NOV-1999; 99WO-US27147.  
XX  
PR 15-DEC-1998; 98US-0112332.  
XX  
PA (PION-) PIONEER HI-BRED INT. INC.  
XX  
PI Mahajan PB;  
XX  
DR WPI. 2000-452026/39.  
XX  
XX N-PSDB; AAA27925.  
XX  
CC The present sequence is that of maize RAD2/FEN-1, as deduced from  
CC a cDNA clone (see AAA27925) derived from maize line W23 tassel  
CC polyA RNA. Rad2/FEN-1 is a structure specific endonuclease which  
CC under certain conditions also acts as an exonuclease. Rad2/FEN-1  
CC can be expressed in transgenic plant cells using conventional  
CC methods. The protein is involved in the regulation of DNA repair  
CC and recombination in plant systems and therefore may be used for  
CC improving gene targeting during further recombinant DNA protocols  
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential  
CC in DNA replication and nucleotide excision and repair reactions.  
CC The exolytic activity is involved in double strand break repair and  
CC end joining. The protein is also useful in strand exchange  
CC reactions during homologous recombination. These functions may be  
CC useful in gene targeting and in the production of male sterile  
CC plants. The efficacy of gene targeting can be improved by the  
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can  
CC be produced by the down regulation of Rad2/FEN-1 expression.  
XX  
SQ Sequence 379 AA;  
  
Query Match 100.0%; Score 379; DB 21; Length 379;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MGIGLTKLADNAPKAMKEQKFESEYFGKRIADVASMISYQFLIVGRTGMETLTNEAGE 60  
DB 1 MGIGLTKLADNAPKAMKEQKFESEYFGKRIADVASMISYQFLIVGRTGMETLTNEAGE 60  
QY 61 VTSHLQGMFNRTRIRLEAGIKPYVYFDGKPPDMKKQELAKRYSKRDATKDLTEAVEVGD 120  
DB 61 VTSHLQGMFNRTRIRLEAGIKPYVYFDGKPPDMKKQELAKRYSKRDATKDLTEAVEVGD 120  
QY 121 KDAIEKLSKRTVAVTROHNEDECKRLRLMGVPVVEAPSEAEACALCINDKVFPAVASED 180  
DB 121 KDAIEKLSKRTVAVTROHNEDECKRLRLMGVPVVEAPSEAEACALCINDKVFPAVASED 180  
QY 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAKVLELELTMDFIDLCILCGDYCDYSIK 240  
DB 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAKVLELELTMDFIDLCILCGDYCDYSIK 240  
QY 241 GIGGOTALKLIRHGSIESILENLNDRQIPEDMPIQEARRLFKEPNVTLDIPELKMTA 300  
DB 241 GIGGOTALKLIRHGSIESILENLNDRQIPEDMPIQEARRLFKEPNVTLDIPELKMTA 300  
QY 301 PDEGLISFLVKDNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
DB 301 PDEGLISFLVKDNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360  
QY 361 DKTSKAANKKTKAGGKK 379  
DB 361 DKTSKAANKKTKAGGKK 379  
  
RESULT 3  
AA95308

```
ID AAY95308 standard; Protein; 379 AA.
XX
XX AAY95308;
AC
XX 12-SEP-2000 (first entry)
DT
XX
XX Maize Rad2/FEN-1 protein.
DE
XX
XX Maize: Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting.
KM
XX
XX Zea mays.
OS
XX WO200036109-A1.
XX
XX 22-JUN-2000.
XX
XX 16-NOV-1999; 99WO-US27147.
XX
XX 15-DEC-1998; 98US-0112332.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Mahajan PB;
XX
XX WPI: 2000-452026/39.
XX
XX N-PSDB: AAA27924.
XX
XX Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -
XX
XX Example 1; Page 74; 85pp; English.
XX
XX The present sequence is that of maize Rad2/FEN-1, as deduced from
CC a cDNA clone (see AAA27924) derived from maize line B73 seedling
CC tissue RNA. Rad2/FEN-1 is a structure specific endonuclease which
CC under certain conditions also acts as an exonuclease. Rad2/FEN-1
CC can be expressed in transgenic plant cells using conventional
CC methods. The protein is involved in the regulation of DNA repair
CC and recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
XX Sequence 379 AA:
SO
Query Match 84.7%; Score 321; DB 21; Length 379;
Best Local Similarity 100.0%; Pred. No. 9.5e-303;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGKIGLTKLADNAPKAMKQKFESEYFGKRIADVASMSTIYOFLLVGRGTMETLTNAGE 60
DB 1 MGKIGLTKLADNAPKAMKQKFESEYFGKRIADVASMSTIYOFLLVGRGTMETLTNAGE 60
QY 61 VTSHLQGMFNRTIRLEAGJIKPYVYFDGKPPDKMKQELAKRSKRDATKDLFEAVVGD 120
DB 61 VTSHLQGMFNRTIRLEAGJIKPYVYFDGKPPDKMKQELAKRSKRDATKDLFEAVVGD 120
QY 121 KDAIEKLSKRTVAVTROHNECKRLLRLMGVPVVEADSEAEACALCINDKYFAVASED 180
DB 121 KDAIEKLSKRTVAVTROHNECKRLLRLMGVPVVEADSEAEACALCINDKYFAVASED 180
QY 181 MDSITFEAPRFLRLMDPSSSKKIPVMEFDVAKYLELELMDQRTIDICICGCGYCSIK 240
DB 181 MDSITFEAPRFLRLMDPSSSKKIPVMEFDVAKYLELELMDQRTIDICICGCGYCSIK 240
```

```
QY 241 GIGGOTALKIRHSGSIESILENLNKRROYIPEDMPYQEARRLKEPNVTLIDPELKWTA 300
DB 241 GIGGOTALKIRHSGSIESILENLNKRROYIPEDMPYQEARRLKEPNVTLIDPELKWTA 300
QY 301 PDEGLISFLVKDNGFNEDRV 321
DB 301 PDEGLISFLVKDNGFNEDRV 321
RESULT 4
AAY95310
ID AAY95310 standard; Protein; 379 AA.
XX
XX AAY95310;
AC
XX 12-SEP-2000 (first entry)
DT
XX
XX Maize Rad2/FEN-1 protein.
DE
XX
XX Maize: Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting.
KM
XX
XX Zea mays.
OS
XX WO200036109-A1.
XX
XX 22-JUN-2000.
XX
XX 16-NOV-1999; 99WO-US27147.
XX
XX 15-DEC-1998; 98US-0112332.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Mahajan PB;
XX
XX WPI: 2000-452026/39.
XX
XX N-PSDB: AAA27926.
XX
XX Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -
XX
XX Example 1; Page 79-80; 85pp; English.
XX
XX The present sequence is that of maize Rad2/FEN-1, as deduced from
CC a cDNA clone (see AAA27926) derived from maize line B73 endosperm
CC RNA. Rad2/FEN-1 is a structure specific endonuclease which
CC under certain conditions also acts as an exonuclease. Rad2/FEN-1
CC can be expressed in transgenic plant cells using conventional
CC methods. The protein is involved in the regulation of DNA repair
CC and recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
XX Sequence 379 AA:
SO
Query Match 73.4%; Score 278; DB 21; Length 379;
Best Local Similarity 99.7%; Pred. No. 4.8e-261;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGKIGLTKLADNAPKAMKQKFESEYFGKRIADVASMSTIYOFLLVGRGTMETLTNAGE 60
DB 1 MGKIGLTKLADNAPKAMKQKFESEYFGKRIADVASMSTIYOFLLVGRGTMETLTNAGE 60
```

QY 61 VTSHLGMEFNTIRLEAGIKPVYVFDGKPPDMKKOELAKRYSRDDATKDLPEAVEVD 120  
 DB 61 VTSHLGMEFNTIRLEAGIKPVYVFDGKPPDMKKOELAKRYSRDDATKDLPEAVEVD 120  
 QY 121 KDAIEKLSKRTVKTQTHNEDCKRLRLMGVPEVPAESEAECALCINDKFAVAVED 180  
 DB 121 KDAIEKLSKRTVKTQTHNEDCKRLRLMGVPEVPAESEAECALCINDKFAVAVED 180  
 QY 181 MDLTFEGAPRFLRLMDPSSSKKIPVMEEDVAKVLEELFLMDQFTDLCICGCDYCSIK 240  
 DB 181 KDSLTFEGAPRFLRLMDPSSSKKIPVMEEDVAKVLEELFLMDQFTDLCICGCDYCSIK 240  
 QY 241 GIGGOTALKLRHSGSIESILENKNDRYQIPEDMPYQEARRLFKEPNVTLDIPELKMTA 300  
 DB 241 GIGGOTALKLRHSGSIESILENKNDRYQIPEDMPYQEARRLFKEPNVTLDIPELKMTA 300  
 QY 301 PDEGLISFLVKDNGFEDRYTAKIEKIKSAKNSOGRLSEFFKPTATTSAPLKRKETS 360  
 DB 301 PDEGLISFLVKDNGFEDRYTAKIEKIKSAKNSOGRLSEFFKPTATTSAPLKRKETS 360  
 QY 361 DKTSAANAANKTKRAGKK 379  
 DB 361 DKTSAANAANKTKRAGKK 379

RESULT 5  
 AAB96644  
 ID AAB96644 standard; Protein; 343 AA.

AC AAB96644;  
 DT 29-OCT-2001 (first entry)  
 DE Putative P. abyssi nuclease.  
 XX  
 KW Hyperthermophilic archaeon; hyperthermophilic protein.  
 XX  
 OS Pyrococcus abyssi.  
 XX  
 PN FR2792651-A1.  
 XX  
 PD 27-OCT-2000.  
 XX  
 PE 21-APR-1999; 99FR-0005034.  
 XX  
 PR 21-APR-1999; 99FR-0005034.  
 XX  
 PA (CNRS) CNRS CENT NAT RECH SCI.  
 (IFRE-) IFREMER INST FR RECH EXPL MER.  
 XX  
 PI Querterleu J, Thierry JC, Prieur D, Dietrich J, Lecompte O;  
 XX  
 DR WPI: 2001-126236/14.

PT New nucleotide sequences isolated from Pyrococcus abyssi encode  
 PT proteins useful in industry -  
 XX  
 XX

PS Claim 7: Pages 1394-1395; 1657pp; French.

CC The present invention relates to the genomic sequence of Pyrococcus  
 CC abyssi (see AAB86431 and AAB41223-7) and P. abyssi proteins. P. abyssi is  
 CC a hyperthermophilic archaeon, which is isolated from deep-sea  
 CC hydrothermal vents. The present sequence is one such P. abyssi protein.  
 CC The proteins of the present invention have various potential industrial  
 CC uses, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade.  
 CC Note: This patent is in the same patent family as WO200065062, which  
 CC contains additional sequences as shown in AAB99132-AAB99143,  
 CC AAB75903-AAH75920 and AAG66436.  
 CC  
 XX  
 SQ Sequence 343 AA;

Query Match 4.0%; Score 15; DB 22; Length 343;  
 Best Local Similarity 100.0%; Pred. No. 5,4e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 77 EAGIKPVYFDGKPP 91  
 DB 70 EAGIKPVYFDGKPP 84

RESULT 6  
 ABB63960  
 ID ABB63960 standard; Protein; 385 AA.

AC ABB63960;  
 DT 26-MAR-2002 (first entry)  
 DE Drosophila melanogaster polypeptide SEQ ID NO 18672.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI: 2001-656860/75.  
 DR N-PSDB; ABL08063.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure: SEQ ID NO 18672; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 385 AA;

Query Match 3.7%; Score 14; DB 22; Length 385;  
 Best Local Similarity 100.0%; Pred. No. 5,6e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVYFDGKPPD 92  
 DB 78 GIKPVYFDGKPPD 91

RESULT 7  
 AAW92505  
 ID AAW92505 standard; Protein; 377 AA.



```

XX AC AAW92505;
XX XX
DT 23-APR-1999 (first entry)
DE Mouse FEN-1 protein.
XX FEN-1; mouse; flap endonuclease; detection; diagnosis; carcinogen;
XX neoplasma; antineoplastic agent; cleavage.
OS Mus sp.
XX US5874283-A.
XX 23-FEB-1999.
XX 30-MAY-1995; 95US-0455968.
XX 30-MAY-1995; 95US-0455968.
XX PR 30-MAY-1995; 95US-0455968.
XX PA (HARR/) HARRINGTON J J.
XX PA (HSIE/) HSIEH C.
XX PA (LIEB/) LIEBER M R.
XX PI Harrington JJ, Hsieh C, Lieber MR;
DR WPI: 1999-179985/15.
DR N-PSDB; AAX02108.
XX PT DNA encoding flap endonuclease polypeptides - useful for producing
XX PT e.g. recombinant polypeptides
XX PS Claim 1: Fig 2A; 58pp; English.
XX CC This sequence represents a mouse FEN-1 (flap endonuclease) protein. This
XX CC protein can be used in methods for detecting a pathological condition in
XX CC a patient, for diagnostic purposes, for screening for antineoplastic
XX CC agents and carcinogens, for diagnostic staging of neoplasia, for
XX CC producing recombinant flap endonuclease for use as research or
XX CC diagnostic reagents, for producing antibodies reactive with the novel
XX CC polypeptides, for producing transgenic nonhuman animals expressing the
XX CC novel polypeptides encoded by a transgene. The invention also provides
XX CC novel molecular cloning techniques and reagents involving cleavage of
XX CC a flap or nick with a flap endonuclease.
XX SQ Sequence 377 AA:
Query Match 3.4%; Score 13; DB 20; Length 377;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 79 GIKPVYVFDGKRP 91
Db 76 GIKPVYVFDGKRP 88

```

```

PD 23-FEB-1999.
PF 30-MAY-1995; 95US-0455968.
XX 30-MAY-1995; 95US-0455968.
XX PR 30-MAY-1995; 95US-0455968.
XX PA (HARR/) HARRINGTON J J.
XX PA (HSIE/) HSIEH C.
XX PA (LIEB/) LIEBER M R.
XX PI Harrington JJ, Hsieh C, Lieber MR;
XX XX
DR WPI: 1999-179985/15.
DR N-PSDB; AAX02111.
XX PT DNA encoding flap endonuclease polypeptides - useful for producing
XX PT e.g. recombinant polypeptides
XX PS Disclosure; Fig 5A-B; 58pp; English.
XX CC This sequence represents a human FEN-1 (flap endonuclease) protein. This
XX CC protein can be used in methods for detecting a pathological condition in
XX CC a patient, for diagnostic purposes, for screening for antineoplastic
XX CC agents and carcinogens, for diagnostic staging of neoplasia, for
XX CC producing recombinant flap endonuclease for use as research or
XX CC diagnostic reagents, for producing antibodies reactive with the novel
XX CC polypeptides, for producing transgenic nonhuman animals expressing the
XX CC novel polypeptides encoded by a transgene. The invention also provides
XX CC novel molecular cloning techniques and reagents involving cleavage of
XX CC a flap or nick with a flap endonuclease.
XX SQ Sequence 378 AA:
Query Match 3.4%; Score 13; DB 20; Length 378;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 79 GIKPVYVFDGKRP 91
Db 76 GIKPVYVFDGKRP 88

```

```

RESULT 8
AAW92508
ID AAW92508 standard; Protein: 378 AA.
XX AAW92508;
XX 23-APR-1999 (first entry)
DE Human FEN-1 protein.
XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
XX neoplasma; antineoplastic agent; cleavage.
OS Homo sapiens.
XX US5874283-A.
XX PN
XX DR

```

```

RESULT 9
AAW92504
ID AAW92504 standard; Protein: 380 AA.
XX AAW92504;
XX 23-APR-1999 (first entry)
DE Human FEN-1 protein.
XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
XX neoplasma; antineoplastic agent; cleavage.
OS Homo sapiens.
XX US5874283-A.
XX 23-FEB-1999.
XX 30-MAY-1995; 95US-0455968.
XX 30-MAY-1995; 95US-0455968.
XX PR 30-MAY-1995; 95US-0455968.
XX PA (HARR/) HARRINGTON J J.
XX PA (HSIE/) HSIEH C.
XX PA (LIEB/) LIEBER M R.
XX PI Harrington JJ, Hsieh C, Lieber MR;
XX WPI: 1999-179985/15.
XX N-PSDB; AAX02107.
XX DR

```

PT	DNA encoding flap endonuclease polypeptides - useful for producing
XX	e.g. recombinant polypeptides
PS	Claim 1; Fig 1A; 58pp; English.
CC	This sequence represents a human FEN-1 (flap endonuclease) protein. This
CC	protein can be used in methods for detecting a pathological condition in
CC	a patient, for diagnostic purposes, for screening for antineoplastic
CC	agents and carcinogens, for diagnostic staging of neoplasms, for
CC	producing recombinant flap endonuclease for use as research or
CC	diagnostic reagents, for producing antibodies reactive with the novel
CC	polypeptides, for producing transgenic nonhuman animals expressing the
CC	novel polypeptides encoded by a transgene. The invention also provides
CC	a flap or nick with a flap endonuclease.
SQ	Sequence      380 AA;
OY	Query Match                  3.4%; Score 13; DB 20; Length 380; Best Local Similarity    100.0%; Pred. No. 0.00052; Matches    13; Conservative    0; Mismatches    0; Indels    0; Gaps    0;
Db	79 GIKPYVFDGKPP 91       78 GIKPYVFDGKPP 90
RESULT 10	
AAM59950	ID AAM59950 standard; Protein; 325 AA.
AC	AAM59950;
DT	22-DEC-1998 (first entry)
DE	Amino acid sequence of the structure specific nuclease 2.
KM	Tag mutant gene; thermostable; nuclease; mutant; DNA polymerase;
XX	bacteria; protozoa; fungi; RNA virus; hepatitis C virus; HCV.
OS	Synthetic.
PX	WO9623774-A1.
PD	04-JUN-1998.
PF	26-NOV-1997; 97MO-US21783.
PR	02-DEC-1996; 96US-0758314. 29-NOV-1996; 96US-0757653.
PA	(THIR-) THIRD WAVE TECHNOLOGIES INC.
PI	Kaiser MW, Lyamichev VI, Lyamicheva N;
WI	WPI; 1998-322748/28.
PT	Thermostable structure-specific nuclease(s) derived from mutant DNA
PT	polymerase(s) - useful for detecting mutant allele(s) or strains of
PT	microorganisms
XX	Claim 1; Page 328; 472pp; English.
XX	This is the amino acid sequence of a structure specific nuclease
CC	used in the method of the invention. In this process thermostable
CC	structure-specific nucleases are derived from mutant DNA polymerases,
CC	which can be used for detecting mutant alleles or strains of
CC	microorganisms. The structure-specific nucleases can be used in
CC	mixtures, compositions and kits to treat nucleic acid, e.g. for
CC	detection of wild type and mutant alleles of genes, for detection
CC	and/or identification of strains of microorganisms such as bacteria,
CC	fungi, protozoa, especially for detection of RNA viruses such as the

```

CC hepatitis C virus (HCV).
XX
SQ Sequence 325 AA:
Query Match 3.2%; Score 12; DB 19; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 77 EAGIKPVYVFDG 88
XXXXXXXXXXXX
DB 69 EAGIKPVYVFDG 80

RESULT 11
AAW59953
ID AAW59953 standard; Protein; 326 AA.
XX
XX AAW59953;
AC
XX 22-DEC-1998 (first entry)
DT
XX Amino acid sequence of the structure specific nuclease 5.
DE
XX
XX Tag mutant gene; thermostable; nuclease; mutant; DNA polymerase;
KM bacteria; protozoa; fungi; RNA virus; hepatitis C virus; HCV.
XX
XX Synthetic.
OS
XX WO9823774-A1.
PN
XX 04-JUN-1998.
PD
XX
XX 26-NOV-1997; 97WO-0521783.
PF
XX
XX 02-DEC-1996; 96US-0758314.
PR
XX 29-NOV-1996; 96DS-0757653.
PI (THIR-) THIRD WAVE TECHNOLOGIES INC.
PA
XX
XX Kaiser MW, Lyamichev VI, Lyamicheva N;
PI
XX WPI: 1998-322748/28.
DR
XX
XX Thermostable structure-specific nuclease(s) derived from mutant DNA
PT polymerase(s) - useful for detecting mutant allele(s) or strains of
PT microorganisms
XX
PS Claim 1; Pages 331-332; 472pp; English.
XX
XX This is the amino acid sequence of a structure specific nuclease
CC used in the method of the invention. In this process thermostable
CC structure-specific nucleases are derived from mutant DNA polymerases,
CC which can be used for detecting mutant alleles or strains of
CC microorganisms. The structure-specific nucleases can be used in
CC mixtures, compositions and kits to treat nucleic acid, e.g. for
CC detection of wild type and mutant alleles of genes, for detection
CC and/or identification of strains of microorganisms such as bacteria,
CC fungi, protozoa, especially for detection of RNA viruses such as the
CC hepatitis C virus (HCV).
XX
XX
SQ Sequence 326 AA:
Query Match 3.2%; Score 12; DB 19; Length 326;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 77 EAGIKPVYVFDG 88
XXXXXXXXXXXX
DB 70 EAGIKPVYVFDG 81

```

```

ID  AAW59951 standard; Protein: 332 AA.
XX
XX  AAW59951;
AC
XX  22-DEC-1998 (first entry)
DT
XX  Amino acid sequence of the structure specific nuclease 3.
DE
XX
XX  Tag mutant gene: thermostable; nuclease; mutant; DNA polymerase;
KW  bacteria; protozoa; fungi; RNA virus; hepatitis C virus; HCV.
XX
XX  Synthetic.
OS
XX  WO9823774-A1.
PN
XX  04-JUN-1998.
PD
XX  26-NOV-1997; 97WO-US21783.
PF
XX  02-DEC-1996; 96US-0758314.
PR  29-NOV-1996; 96US-0757653.
PA  (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX  Kaiser MW, Lyamichev VI, Lyamicheva N;
PI
XX  WPI: 1998-322748/28.
DR
XX  Thermostable structure-specific nuclease(s) derived from mutant DNA
PT  polymerase(s) - useful for detecting mutant allele(s) or strains of
PT  microorganisms
XX
XX  Claim 1; Pages 329-330; 472pp; English.
PS
XX  This is the amino acid sequence of a structure specific nuclease
CC  used in the method of the invention. In this process thermostable
CC  structure-specific nucleases are derived from mutant DNA polymerases,
CC  which can be used for detecting mutant alleles or strains of
CC  microorganisms. The structure-specific nucleases can be used in
CC  mixtures, compositions and kits to treat nucleic acid, e.g. for
CC  detection of wild type and mutant alleles of genes, for detection
CC  and/or identification of strains of microorganisms such as bacteria,
CC  fungi, protozoa, especially for detection of RNA viruses such as the
CC  hepatitis C virus (HCV).
XX
XX  Sequence 332 AA:
SQ
XX
XX  Query Match 3.2%; Score 12; DB 19; Length 332;
XX  Best Local Similarity 100.0%; Pred. No. 0.0043;
XX  Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  77 EAGIKPVYVFDG 88
DB  70 EAGIKPVYVFDG 81

```

```

PD  31-JUL-1997.
XX
XX  22-JAN-1997; 97WO-US01072.
PF
XX  02-DEC-1996; 96US-0759038.
PR  24-JAN-1996; 96US-0599491.
PR  12-JUL-1996; 96US-0682853.
PR  29-NOV-1996; 96US-0756386.
PR  02-DEC-1996; 96US-0758314.
XX
XX  (THIR-) THIRD WAVE TECHNOLOGIES INC.
PA
XX  Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichev VI;
PI  Olive DM, Prudent JR;
XX
XX  WPI: 1997-393613/36.
DR  N-PSDB: AAT76685.
PT
PT  Thermostable structure-specific nuclease(s) - used for detection and
PT  characterisation of nucleic acid sequences and variations in nucleic
PT  acid sequences
XX
XX  Example 28; Page 283-285; 457pp; English.
PS
XX
XX  This sequence comprises Pyrococcus furiosus (Pfu) FEN-1
CC  endonuclease. Large-scale production of the enzyme was performed
CC  using E. coli host cells transformed with a vector carrying the Pfu
CC  FEN-1 coding sequence (see AAT76685). Pfu FEN-1 is a thermostable
CC  enzyme. It can be used in novel methods for the detection and
CC  characterisation of nucleic acid sequences and variations in
CC  nucleic acid sequences.
XX
XX  Sequence 340 AA:
SQ
XX
XX  Query Match 3.2%; Score 12; DB 18; Length 340;
XX  Best Local Similarity 100.0%; Pred. No. 0.0044;
XX  Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  77 EAGIKPVYVFDG 88
DB  70 EAGIKPVYVFDG 81

```

```

RESULT 13
AAW24216
ID  AAW24216 standard; Protein: 340 AA.
XX
XX  AAW24216;
AC
XX  14-APR-1998 (first entry)
DT
XX  Pyrococcus furiosus FEN-1 endonuclease.
DE
XX
XX  Nucleic acid cleavage; 5' nuclease; DNA cleavage; RNA cleavage;
KW  invader directed cleavage; FEN-1; endonuclease.
XX
XX  Pyrococcus furiosus.
OS
XX  WO9727214-A1.
DR

```

```

RESULT 14
AAW79970
ID  AAW79970 standard; Protein: 340 AA.
XX
XX  AAW79970;
AC
XX  02-FEB-1999 (first entry)
DT
XX  Pyrococcus furiosus FEN-1 endonuclease.
DE
XX
XX  Nucleic acid detection; multiple sequential invasive cleavage;
KW  FEN-1; endonuclease; nuclease.
XX
XX  Pyrococcus furiosus.
OS
XX  WO9842873-A1.
PN
XX  01-OCT-1998.
PD
XX  24-MAR-1998; 98WO-US05809.
PF
XX  24-MAR-1997; 97US-0823516.
PR
XX
XX  (THIR-) THIRD WAVE TECHNOLOGIES INC.
PA  Brow MAD, Hall JG, Kwiatkowski RW, Lyamichev VI;
PI  Mast AL, Vavra SH;
XX
XX  WPI: 1998-557036/47.
DR  N-PSDB: AAV65840.

```

```
XX Detecting target nucleic acid by sequence-specific cleavage of
PT complex with two specific oligonucleotides - used to detect
PT cytomegalovirus DNA
XX
PS Example 28b, Page 316-317; 524pp; English.
XX
CC This is the amino acid sequence of FEN-1 endonuclease of
CC Pyrococcus furiosus (Pfu). FEN-1 DNA (see AAV65840) has been
CC ligated into vector for expression in E. coli cells. The invention
CC relates to means for the detection and characterization of nucleic
CC acid sequences, and variations in nucleic acid cleavage structure. It also
CC relates to methods for forming a nucleic acid cleavage structure on
CC a target sequence and cleaving this structure in a site-specific
CC manner, preferably using a thermostable structure-specific nuclease
CC such as FEN-1. Cleavage of the cleavage structure by the
CC nuclease indicates the presence of specific nucleic acid sequences
CC or specific variants. The invention further relates to methods for
CC the separation of nucleic acid molecules based on charge, methods
CC for the detection of non-target cleavage products via the formation
CC of a complete and activated protein binding region, and methods for
CC the detection of nucleic acid from various viruses (e.g. human
CC cytomegalovirus) in a sample. The method amplifies the detection
CC molecule rather than the target itself, is less subject to
CC contamination than exponential amplification processes, and allows
CC many targets to be analysed in a single reaction.
CC
SQ Sequence 340 AA:
Query Match 3.2%; Score 12; DB 19; Length 340;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 77 EAGIKPYVFDG 88
|||||
Db 70 EAGIKPYVFDG 81
RESULT 15
AAW59940
ID AAW59940 standard; Protein; 340 AA.
XX
AC AAW59940;
XX
DT 21-DEC-1998 (first entry)
XX
DE Amino acid sequence of the Pfu FEN-1 endonuclease.
XX
KW Mja FEN-1 endonuclease; Tag gene; structure-specific nuclease;
KW mutant; DNA polymerase; bacteria; fungi; protozoa; RNA virus;
KW hepatitis C virus; HCV; thermostable.
XX
OS Pyrococcus furiosus.
XX
PN WO9823774-A1.
XX
PD 04-JUN-1998.
XX
PF 26-NOV-1997; 97WO-US21783.
XX
PR 02-DEC-1996; 96US-0758314.
XX
PR 29-NOV-1996; 96US-0757653.
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Kaiser MM, Lyamichev VI, Lyamicheva N;
XX
DR WPI, 1998-322748/28.
XX
DR N-PSDB; AAV53951.
XX
PT Thermostable structure-specific nuclease(s) derived from mutant DNA
PT polymerase(s) - useful for detecting mutant allele(s) or strains of
PT microorganisms
```

```
XX Example 29; Pages 280-281; 472pp; English.
XX
CC This is the amino acid sequence encoding the Pfu FEN-1 endonuclease,
CC used in the method of the invention. In this process thermostable
CC structure-specific nucleases are derived from mutant DNA polymerases,
CC which can be used for detecting mutant alleles or strains of
CC microorganisms. The structure-specific nucleases can be used in
CC mixtures, compositions and kits to treat nucleic acid, e.g. for
CC detection of wild type and mutant alleles of genes, for detection
CC and/or identification of strains of microorganisms such as bacteria,
CC fungi, protozoa, especially for detection of RNA viruses such as the
CC hepatitis C virus (HCV).
XX
SQ Sequence 340 AA:
Query Match 3.2%; Score 12; DB 19; Length 340;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 77 EAGIKPYVFDG 88
|||||
Db 70 EAGIKPYVFDG 81
RESULT 16
AAW92511
ID AAW92511 standard; Protein; 16 AA.
XX
AC AAW92511;
XX
DT 23-APR-1999 (first entry)
XX
DE Human FEN-1 protein epitope #2.
XX
KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage; epitope.
XX
OS Homo sapiens.
XX
PN US5874283-A.
XX
PD 23-FEB-1999.
XX
PF 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
PI Harrington JJ, Hsieh C, Lieber MR;
XX
DR WPI; 1999-179985/15.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing
PT e.g. recombinant polypeptides
XX
PS Disclosure; Column 22; 58pp; English.
XX
XX AAW92410-W92522 represent epitopes from the human FEN-1 (flap
XX endonuclease) protein. This protein can be used in methods for detecting
XX a pathological condition in a patient, for diagnostic purposes, for
XX screening for antineoplastic agents and carcinogens, for diagnostic
XX staging of neoplasia, for producing recombinant flap endonuclease for use
XX as research or diagnostic reagents, for producing antibodies reactive
XX with the novel polypeptides, for producing transgenic nonhuman animals
XX expressing the novel polypeptides encoded by a transgene. The invention
XX also provides novel molecular cloning techniques and reagents involving
XX cleavage of a flap or nick with a flap endonuclease.
XX
SQ Sequence 16 AA:
```

Query Match 2.4%; Score 9; DB 20; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 0.21;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 SMSITQFLI 44  
 |||||||  
 DB 1 SMSITQFLI 9

## RESULT 17

AAW92506  
 ID AAW92506 standard; Protein; 382 AA.

XX AAW92506;

DT 23-APR-1999 (first entry)

DE Yeast FEN-1 protein.

XX FEN-1; Yeast; flap endonuclease; detection; diagnosis; carcinogen;  
 KW neoplasma; antineoplastic agent; cleavage.

OS Saccharomycetes cerevisiae.

XX US5874283-A.

PN 23-FEB-1999.

PD 30-MAY-1995; 95US-0455968.

PR 30-MAY-1995; 95US-0455968.

XX (HARR/) HARRINGTON J J.

PA (HSTE/) HSTEH C.

XX (LIEB/) LIEBER M R.

PI Harrington JJ, Hsieh C, Lieber MR;

XX WPI: 1999-179985/15.

DR N-PSDB: AAX02109.

XX DNA encoding flap endonuclease polypeptides - useful for producing

PT e.g. recombinant polypeptides

XX Disclosure: Fig 3A; 58pp; English.

CC This sequence represents a yeast FEN-1 (flap endonuclease) protein. This  
 CC protein is used in a method to isolate novel human FEN-1 proteins for  
 CC detecting a pathological condition in a patient, for diagnostic purposes,  
 CC for screening for antineoplastic agents and carcinogens, for diagnostic  
 CC staging of neoplasia, for producing recombinant flap endonuclease for use  
 CC as research or diagnostic reagents, for producing antibodies reactive  
 CC with the novel polypeptides, for producing transgenic nonhuman animals  
 CC expressing the novel polypeptides encoded by a transgene. The invention  
 CC also provides novel molecular cloning techniques and reagents involving  
 CC cleavage of a flap or nick with a flap endonuclease.

XX SQ Sequence 382 AA;

Query Match 2.4%; Score 9; DB 20; Length 382;

Best Local Similarity 100.0%; Pred. No. 4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 84 YVFDGKPPD 92  
 |||||||

DB 84 YVFDGKPPD 92

## RESULT 18

AA10723

ID AA10723 standard; Protein; 145 AA.  
 XX

AC AA10723;

DT 22-APR-1991 (first entry)

XX (Lys79)IL-3.

XX Interleukin-3; mutant; antagonist; cytopenia; immunodeficiency;

KW immunosuppression; charge reversal.

XX Key location/Qualifiers

FT Peptide 1..12

FT Mutation /label= signal sequence

XX /label= Lys -> Glu

PN EP413383-A.

PD 20-FEB-1991.

XX 02-AUG-1990; 90EP-0202117.

PR 14-SEP-1989; 89EP-0202331.

PR 14-AUG-1989; 89EP-0202082.

XX 02-AUG-1990; 90EP-0202117.

PA (KONN ) GIST-BROCADES NV.

XX Dorsiers LCI, Vanleen RW;

PI WPI: 1991-052685/08.

DR New interleukin-3 mutants - having deletions covering complete

XX coding sequence while retaining biological activity, have

PT antagonistic effect.

XX Claim 6; Page 18; 26pp; English.

PS The mutant was prepd. by recombinant DNA techniques using mutagenic

XX primers to change the codons of cDNA encoding IL-3. The peptide

CC pref. also has a deletion at the N-terminus (AAs 1-14) and/or the

CC C-terminus (AAs 116-133,120-130 and 130-133). The peptide has

CC antagonistic activity and can be used for therapeutic and diagnostic

CC purposes.

XX See also AA10711-R10731 and AA11051-R11055.

XX SQ Sequence 145 AA;

Query Match 2.1%; Score 8; DB 12; Length 145;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 257 IESILENL 264  
 |||||||

DB 86 IESILENL 93

## RESULT 19

ABG08130  
 ID ABG08130 standard; Protein; 338 AA.

XX ABG08130;

DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #8121.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-0508631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI: 2001-639362/73.  
 DR N-PSDB; AAS72317.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX  
 PS Claim 20; SEQ ID NO 38489; 103bp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 338 AA;  
 XX  
 Query Match 2.1%; Score 8; DB 22; Length 338;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 123 ALEKLSKR 130  
 |||||  
 Db 273 ALEKLSKR 280  
 XX  
 RESULT 20  
 AAY03778  
 ID AAY03778 standard; Protein; 343 AA.  
 XX  
 AC AAY03778;  
 XX  
 DT 14-JUN-1999 (first entry)  
 XX  
 DE Pyrococcus heat resistant Flap endonuclease.  
 XX  
 KM Flap endonuclease; thermophile; thermostable; heat resistant; enzyme;  
 KM genetic engineering; recombination; gene shuffling.  
 XX  
 OS Pyrococcus sp.  
 XX  
 PN JP11075849-A.  
 XX  
 PD 23-MAR-1999.

PF 04-SEP-1997; 97JP-0239440.  
 XX  
 PR 04-SEP-1997; 97JP-0239440.  
 XX  
 PA (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.  
 XX  
 DR WPI: 1999-257697/22.  
 DR N-PSDB; AAX31850.  
 XX  
 PT Heat resistant Flap endonuclease - derived from an extreme  
 PT thermophile of Pyrococcus genus  
 XX  
 PS Claim 2; Page 5-6; 8pp; Japanese.  
 XX  
 CC This represents a heat resistant Flap endonuclease derived from an  
 CC extreme thermophile of Pyrococcus genus. Transformants containing a  
 CC vector comprising the DNA are cultured for the recombinant expression of  
 CC the thermostable Flap endonuclease. The enzyme is thermostable (having  
 CC an optimum temperature of at least of 75 deg C) and can be used in  
 CC standard genetic engineering techniques, especially for the development  
 CC of new methods for efficiently carrying out homologous recombination and  
 CC gene shuffling.  
 XX  
 SQ Sequence 343 AA;  
 XX  
 Query Match 2.1%; Score 8; DB 20; Length 343;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 84 YVFDGKRP 91  
 |||||  
 Db 77 YVFDGKRP 84  
 XX  
 RESULT 21  
 AAP70501  
 ID AAP70501 standard; Protein; 373 AA.  
 XX  
 AC AAP70501;  
 XX  
 DT 13-MAY-1991 (first entry)  
 XX  
 DE Chinese hamster glutamine synthetase gene product.  
 XX  
 KW DHFR/MTX.  
 XX  
 OS Cricetus griseus.  
 XX  
 PN W08704462-A.  
 XX  
 PD 30-JUL-1987.  
 XX  
 PF 23-JAN-1987; 87WO-G000039.  
 XX  
 PR 23-JAN-1986; 86GB-0001597.  
 XX  
 PA (CELL-) CELTECH LTD.  
 PA (UNIU ) UNIV OF GLASGOW.  
 PA (WILS/) WILSON RH.  
 XX  
 PI Wilson RH, Bebbington CR;  
 XX  
 DR WPI: 1987-221263/31.  
 DR N-PSDB; AAN70819.  
 XX  
 PT Recombinant DNA which encodes glutamine synthetase - used esp. in  
 PT co-amplification of non-selected genes and in transforming host  
 PT cell lines.  
 XX  
 PS Claim 6; Fig 2a-e; 54pp; English.  
 XX  
 CC Sequence encoding the product may be used as a hybridisation probe in  
 CC the isolation of the GS gene in other species, and as a diagnostic tool

CC in the detection of diseases altering GS expression. A second use is  
 CC as a selectable marker in recombinant DNA technology eg. co-amplification  
 CC selection, or transformation to glutamine independence.

XX  
 SQ Sequence 373 AA:

Query Match 2.1%; Score 8; DB 8; Length 373;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 AIEKLSKR 130  
 |||||  
 Db 273 AIEKLSKR 280

# RESULT 22

ABB59358 ID ABB59358 standard; Protein; 399 AA.

XX ABB59358;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 4866.

KW Drosophila; developmental biology; cell signalling; insecticide;  
 pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL03461.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -

PS Disclosure; SEQ ID NO 4866; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 399 AA:

Query Match 2.1%; Score 8; DB 22; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 AIEKLSKR 130  
 |||||  
 Db 309 AIEKLSKR 316

# RESULT 23

ABB65740 ID ABB65740 standard; Protein; 399 AA.

XX ABB65740;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 24012.

KW Drosophila; developmental biology; cell signalling; insecticide;  
 pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL09843.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -

PS Disclosure; SEQ ID NO 24012; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 399 AA:

Query Match 2.1%; Score 8; DB 22; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 AIEKLSKR 130  
 |||||  
 Db 309 AIEKLSKR 316

# RESULT 24

AAV95714 ID AAV95714 standard; Protein; 515 AA.

XX AAV95714;

XX 25-OCT-2000 (first entry)

DE Cosmid cHRIM5 encoded protein P20-9r.

XX Cosmid cHRIM5; nematocidae; nematode; biological control agent;

XX	transgenic plant; helminthiasis; P20-9r.
KM	
OS	Xenorhabdus bovienii.
PN	WO20042855-A1.
XX	
XX	27-JUL-2000.
PD	
XX	24-JAN-2000; 2000WO-GB00219.
PF	
XX	22-JAN-1999; 99GB-0001499.
PR	
XX	(HORT-) HORTICULTURE RES INT.
PA	
XX	Morgan JAW, Jarrett P, Ellis D, Ousley MA;
PI	WPI; 2000-499157/44.
XX	N-PSDB; AAA50029.
DR	
XX	Novel composition used to control parasitic nematodes, especially in
PT	plants such as maize, cotton, soya, and rice, comprises a bacterium
PT	which is a symbiont of an entomopathogenic nematode -
XX	
XX	Example 6; Page 44; 74pp; English.
PS	
XX	The present sequence is that of protein P20-9r encoded by an open
CC	reading frame identified in cosmid cHR15 (see AAA50029). cHR15 was
CC	obtained by ligating Xenorhabdus bovienii strain 173 (NCIMB 40986)
CC	Sau3A-digested DNA fragments into the BamHI site of the Stratagene
CC	cosmid vector Supercos1, packaging into Escherichia coli XL Blue 1,
CC	and screening for nematocidal activity against Caenorhabditis elegans.
CC	Analysis of the DNA indicated a number of open reading frames for
CC	which the corresponding protein sequences were determined (see
CC	AAV5685-1957575). Nematodes can be controlled through the use of
CC	bacteria associated symbiotically with an entomopathogenic nematode.
CC	Such bacteria include Xenorhabdus and Photorhabdus spp. such as X.
CC	bovienii strain 173. The symbiont bacteria, an engineered
CC	bacterium, or a nematocidal protein obtained from such bacteria,
CC	can be used to control helminthiasis in a human or domesticated
CC	animal or for the control of plant pathogen nematodes. Also
CC	claimed are vectors for expressing nematocidal proteins in host
CC	cells, and transgenic plants.
XX	
SO	Sequence 515 AA:
Query Match	2.1%; Score 8; DB 21; Length 515;
Best Local Similarity	100.0%; Pred. No. 49;
Matches 8; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Y	144 RLLRLMGV 151
Db	480 RLLRLMGV 487
RESULT 25	
ID	AAV58997
XX	AAV58997 standard; Protein: 667 AA.
AC	
XX	AAV58997;
XX	
DT	23-MAY-2000 (first entry)
XX	
DE	Soybean sulfite reductase.
XX	
KM	Soybean; sulfite reductase; sulfate assimilation; transgenic plant.
XX	
OS	Glycine max.
XX	
PN	WO200004162-A2.
XX	
PD	27-JAN-2000.
XX	
PF	13-JUL-1999; 99WO-US15812.



FT Region 736..739  
 FT /note="consensus sequence for N-glycosylation"  
 XX  
 PN W09953067-A2.  
 XX  
 PD 21-OCT-1999.  
 XX  
 PF 13-APR-1999; 99WO-CA00293.  
 XX  
 PR 13-APR-1998; 98US-0059090.  
 XX  
 PA (MTAC ) CANADA MIN AGRIC & AGRI-FOOD CANADA.  
 XX  
 PI Mikl B, Gilzen M, Miller S, Bowman L, Batchelor A, Hu M;  
 PI Boullier K;  
 DR WPI: 1999-611304/52.  
 DR N-PSDB; AA229905.  
 XX  
 PT Novel promoter sequences and genes useful for inducing expression of  
 PT genes in plant seed coats -  
 XX  
 PS Example: Fig 23(B); 155pp; English.  
 XX  
 CC The present sequence represents a SC20 protein. The promoter of  
 CC this gene is differentially expressed in seed coat tissues,  
 CC specifically within the outer integument, the inner integument, the  
 CC thick walled parenchyma, the thin walled parenchyma, the endothelium,  
 CC the hourglass cells, the palisade, the stellate parenchyma, or the  
 CC membranous endocarp associated with the seed coat. The seed-coat  
 CC promoters may be used to express proteins of interest in seed coat  
 CC tissues. Genes of interest include but are not restricted to herbicide  
 CC resistance genes, genes encoding viral coat proteins, or genes encoding  
 CC proteins conferring biological control of pests or pathogens, e.g. a Bt  
 CC toxin. Other genes that may be expressed include proteins that alter the  
 CC taste of the seed and/or affect the nutritive value of the seed.  
 CC  
 XX  
 SQ Sequence 770 AA;  
 Query Match 2.1%; Score 8; DB 20; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 152 PYVEAPSE 159  
 |||||  
 DB 707 PYVEAPSE 714  
 RESULT 27  
 AAM96822  
 ID AAM96822 standard; Peptide; 14 AA.  
 AC  
 XX AAM96822;  
 XX  
 DT 24-JAN-2002 (first entry)  
 XX  
 DE Human peptide #97 encoded by a SNP oligonucleotide.  
 XX  
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;  
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
 KW multifactorial disease; autoimmune disease; infection;  
 KW nervous system disease.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 PN W0200147944-A2.  
 XX  
 PD 05-JUL-2001.  
 XX

PF 28-DEC-2000; 2000WO-US35498.  
 XX  
 XX 28-DEC-1999; 99US-0173419.  
 PR 27-DEC-2000; 2000US-0173419.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkets RA, Leach M;  
 DR WPI: 2001-465210/50.  
 XX  
 PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 PT oncogenes and histones, useful for diagnosing and treating, e.g.  
 PT cancer, autoimmune diseases and infections -  
 XX  
 PS Disclosure: Page 3689; 4143pp; English.  
 XX  
 CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)  
 CC encoding polymorphic variants of proteins related to amylases, amyloid  
 CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,  
 CC polymerase, oncogenes, histones, kinases, colony stimulating factors,  
 CC complement related proteins, cytochromes, kinesins, cytokines,  
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.  
 CC The present sequence is a peptide encoded by one such oligonucleotide.  
 CC The oligonucleotides and the peptides encoded by them may be used in the  
 CC prevention, diagnosis and treatment of diseases associated with  
 CC inappropriate expression of the proteins listed above. Disorders that may  
 CC be prevented, diagnosed and/or treated include multifactorial diseases  
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus  
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,  
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous  
 CC system and an infection of pathogenic organisms.  
 CC  
 XX  
 SQ Sequence 14 AA;  
 Query Match 1.8%; Score 7; DB 22; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 179 EDMDSLTL 185  
 |||||  
 DB 4 EDMDSLTL 10  
 RESULT 28  
 AAM92515  
 ID AAM92515 standard; Protein; 25 AA.  
 AC  
 XX AAM92515;  
 XX  
 DT 23-APR-1999 (first entry)  
 XX  
 DE Human FEN-1 protein epitope #6.  
 XX  
 KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;  
 KW neoplasia; antineoplastic agent; cleavage; epitope.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 PN US5874283-A.  
 PD 23-FEB-1999.  
 XX  
 PF 30-MAY-1995; 95US-0455968.  
 XX  
 PR 30-MAY-1995; 95US-0455968.  
 XX  
 PA (HARR/) HARRINGTON J J.  
 PA (HSIEH/) HSIEH C.  
 PA (LIEBER/) LIEBER M R.  
 XX  
 PI Harrington JJ, Hsieh C, Lieber MR;

XX WPI; 1999-179985/15.  
XX DNA encoding flap endonuclease polypeptides - useful for producing  
PT e.g. recombinant polypeptides  
XX  
PS Disclosure; Column 22; 58pp; English.  
XX  
CC AAM92410-W92522 represent epitopes from the human FEN-1 (flap  
CC endonuclease) protein. This protein can be used in methods for detecting  
CC a pathological condition in a patient, for diagnostic purposes, for  
CC screening for antineoplastic agents and carcinogens, for diagnostic  
CC staging of neoplasia, for producing recombinant flap endonuclease for use  
CC as research or diagnostic reagents, for producing antibodies reactive  
CC with the novel polypeptides, for producing transgenic nonhuman animals  
CC expressing the novel polypeptides encoded by a transgene. The invention  
CC also provides novel molecular cloning techniques and reagents involving  
CC cleavage of a flap or nick with a flap endonuclease.  
CC  
XX  
SQ Sequence 25 AA;  
  
Query Match 1 8%; Score 7; DB 20; Length 25;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 156 APSEAEA 162  
Db 12 APSEAEA 18  
|||||  
12 APSEAEA 18  
  
RESULT 29  
AAM90755  
ID AAM90755 standard; Protein; 49 AA.  
AC AAM90755;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen SEQ ID NO:18348.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis.  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225457.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234927.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 12-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 06-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Barash SC, Ruben SM;  
 PI WPI: 2001-483426/52.  
 DR N-PSDB; AAK63536.  
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 XX  
 PS Claim 11: SEQ ID NO 18348; 3071pp + Sequence Listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/haematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/haematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention.  
 CC  
 XX

SO Sequence 49 AA;  
 Query Match 1.8%; Score 7; DB 22; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 346 PTAITSA 352  
 |||||  
 Db 31 PTAITSA 37  
 RESULT 30  
 AAB34731  
 ID AAB34731 standard; Protein; 61 AA.  
 XX AAB34731;  
 AC AAB34731;  
 DT 26-JAN-2001 (first entry)  
 XX  
 DE Human secreted protein encoded by DNA clone vq13 1.  
 XX  
 KW Secreted protein; human; autoimmune disorder; multiple sclerosis; ulcer;  
 KW systemic lupus erythematosus; rheumatoid arthritis; anaemia; stroke;  
 KW haematopoiesis regulation; tissue growth; wound healing; haemophilia;  
 KW Alzheimer's disease; Parkinson's disease; Shy-drager syndrome; cancer;  
 KW contraceptive; infection; growth inhibition; hyperproliferative disorder;  
 KW poriasis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055375-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 17-MAR-2000; 2000MO-US07285.  
 XX  
 PR 17-MAR-1999; 99US-0124808.  
 PR 17-MAR-1999; 99US-0124916.  
 PR 17-AUG-1999; 99US-0149639.  
 PR 01-OCT-1999; 99US-0157247.  
 PR 29-NOV-1999; 99US-0167824.  
 PR 15-FEB-2000; 2000US-0182711.  
 XX  
 PA (ALPH-) ALPHAGENE INC.  
 XX  
 PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;  
 XX  
 DR WPI: 2000-638211/61.  
 DR N-PSDB; AAC59832.  
 XX  
 PT Novel proteins and polypeptides useful for the treatment of e.g  
 PT multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,  
 PT cancer, Alzheimer's disease, Parkinson's disease, stroke, anemia and  
 PT ulcers -  
 XX  
 PS Claim 98; Page 444-445; 493pp; English.  
 XX  
 CC This invention relates to 59 human secreted proteins and the nucleotide  
 CC sequences encoding them. Sequences AAC59788-C59846 and AAB34687-834745  
 CC represent the proteins and their encoding nucleotide sequences, and  
 CC sequences AAB34746-834771 represent fragments of the proteins. Probes  
 CC for the DNA sequences are represented by sequences AAC59847-C599596. The  
 CC proteins exhibit neuroprotective, dermatological, immunosuppressive,  
 CC antiinflammatory, antiaemic, nootropic, antiparkinsonian,  
 CC cerebroprotective, haemostatic, vulnerary, cytoparctic, antipruritic,  
 CC antibacterial, virucide, and fungicide activity. The proteins and  
 CC nucleotide sequences are useful as nutritional sources or supplements  
 CC and in research. The proteins are useful for treating immune deficiency  
 CC and disorders, which may be genetic or resulting from infections,  
 CC autoimmune disorders such as multiple sclerosis, systemic lupus  
 CC erythematosus, rheumatoid arthritis, and for treating myeloid or lymphoid  
 CC cell deficiencies such as anaemias by regulating haematopoiesis. The  
 CC proteins are also useful in compositions for bone, cartilage, tendon,  
 CC



CC nucleic acid or the presence or amount of expression of a secreted  
CC protein. The sequences are used to prevent, treat or ameliorate a medical  
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. The antibodies to the polypeptides can also be used in  
CC alleviating symptoms associated with disorders and in diagnostic  
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent  
CC assays (ELISA). The disorders include autoimmune diseases e.g. rheumatoid  
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or  
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular  
CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders  
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
CC fungi and ocular disorders e.g. corneal infection. The peptides can also  
CC be used to aid wound healing and epithelial cell proliferation, to help  
CC prevent skin ageing due to sunburn, to maintain organs before  
CC transplantation, to regenerate tissues, in chemotaxis and as a food  
CC additive or preservative to alter storage capabilities.  
CC  
XX  
SQ Sequence 72 AA;

Query Match 1.8%; Score 7; DB 22; Length 72;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 276 PYQEARR 282  
| | | | | | |  
Db 43 PYQEARR 49

RESULT 33  
AAG05446  
ID AAG05446 standard; Protein: 73 AA.  
XX  
XX AAG05446;  
AC  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1855.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
PN EP1033405-A2.  
XX  
XX 06-SEP-2000.  
PD  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158362.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 7; DB 21; Length 73;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 KDAIEKL 127  
Db 65 KDAIEKL 71

RESULT 34  
AAG07581  
ID AAG07581 standard; Protein: 73 AA.  
XX AAG07581;  
AC AAG07581;  
XX 17-OCT-2000 (first entry)  
DT  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4789.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hydridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX  
PN EPI033405-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 09-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134376.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149802.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159658.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 7; DB 21; Length 73;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KDAIEKL 127  
Db 65 KDAIEKL 71

RESULT 35  
AAG08148  
ID AAG08148 standard; Protein: 73 AA.

AC AAG08148;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 5572.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134376.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139482.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
XX 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145195.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.



PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147435.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151080.  
PR 27-AUG-1999; 99US-0151086.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161932.  
PR 28-OCT-1999; 99US-0161933.  
PR 29-OCT-1999; 99US-0162142.

Query Match 1.88; Score 7; DB 21; Length 73;  
Best Local Similarity 100.0%; Pred. No. 75;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 121 KDAIEKL 127  
1111111  
Db 65 KDAIEKL 71  
RESULT .36  
AAG08766  
ID AAG08766 standard: Protein; 73 AA.  
XX  
AC AAG08766;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6433.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149358.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149920.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158322.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161406.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161820.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Query Match

Best Local Similarity 1.00.0%; Score 7; DB 21; Length 73;  
Pred. No. 75;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KDAIEKL 127  
|||||

Db 65 KDAIEKL 71

RESULT 37  
AAG36961  
ID AAG36961 standard; Protein: 73 AA.  
XX AAG36961;  
AC  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SRO ID NO: 45372.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
PD  
XX 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-030439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145099.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 9905-0149929.  
PR 23-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150566.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151065.  
PR 27-AUG-1999; 9905-0151066.  
PR 30-AUG-1999; 9905-0151080.  
PR 31-AUG-1999; 9905-0151303.  
PR 01-SEP-1999; 9905-0151438.  
PR 07-SEP-1999; 9905-0152363.  
PR 10-SEP-1999; 9905-0153070.  
PR 13-SEP-1999; 9905-0153758.  
PR 15-SEP-1999; 9905-0154018.  
PR 16-SEP-1999; 9905-0154039.  
PR 20-SEP-1999; 9905-0154779.  
PR 22-SEP-1999; 9905-0155139.  
PR 23-SEP-1999; 9905-0155486.  
PR 24-SEP-1999; 9905-0155659.  
PR 28-SEP-1999; 9905-0156458.  
PR 29-SEP-1999; 9905-0156596.  
PR 04-OCT-1999; 9905-0157117.  
PR 05-OCT-1999; 9905-0157753.  
PR 06-OCT-1999; 9905-0157865.  
PR 07-OCT-1999; 9905-0158029.  
PR 08-OCT-1999; 9905-0158232.  
PR 12-OCT-1999; 9905-0158369.  
PR 13-OCT-1999; 9905-0159293.  
PR 13-OCT-1999; 9905-0159294.  
PR 13-OCT-1999; 9905-0159299.  
PR 14-OCT-1999; 9905-0159329.  
PR 14-OCT-1999; 9905-0159330.  
PR 14-OCT-1999; 9905-0159331.  
PR 14-OCT-1999; 9905-0159637.  
PR 14-OCT-1999; 9905-0159688.  
PR 18-OCT-1999; 9905-0159584.  
PR 21-OCT-1999; 9905-0160741.  
PR 21-OCT-1999; 9905-0160767.  
PR 21-OCT-1999; 9905-0160768.  
PR 21-OCT-1999; 9905-0160770.  
PR 21-OCT-1999; 9905-0160814.  
PR 21-OCT-1999; 9905-0160815.  
PR 22-OCT-1999; 9905-0160980.  
PR 22-OCT-1999; 9905-0160981.  
PR 22-OCT-1999; 9905-0160989.  
PR 25-OCT-1999; 9905-0161404.  
PR 25-OCT-1999; 9905-0161405.  
PR 25-OCT-1999; 9905-0161406.  
PR 26-OCT-1999; 9905-0161359.  
PR 26-OCT-1999; 9905-0161360.  
PR 26-OCT-1999; 9905-0161361.  
PR 28-OCT-1999; 9905-0161920.  
PR 28-OCT-1999; 9905-0161992.  
PR 28-OCT-1999; 9905-0161995.  
PR 29-OCT-1999; 9905-0162142.

Query Match 1.8%; Score 7; DB 21; Length 73;  
Best Local Similarity 100.0%; Pred. No. 75;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 KDAIEKL 127  
Db 65 KDAIEKL 71

RESULT 38  
AAG52931  
ID AAG52931 standard; Protein; 73 AA.  
XX AAG52931;  
AC  
XX  
DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 67337.  
DE  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
PN EPI033405-A2.  
PD  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 9905-0121825.  
PR 05-MAR-1999; 9905-0123180.  
PR 09-MAR-1999; 9905-0123548.  
PR 23-MAR-1999; 9905-0125788.  
PR 25-MAR-1999; 9905-0126264.  
PR 29-MAR-1999; 9905-0126785.  
PR 01-APR-1999; 9905-0127462.  
PR 06-APR-1999; 9905-0128234.  
PR 08-APR-1999; 9905-0128714.  
PR 16-APR-1999; 9905-0129845.  
PR 19-APR-1999; 9905-0130077.  
PR 21-APR-1999; 9905-0130449.  
PR 23-APR-1999; 9905-0130510.  
PR 23-APR-1999; 9905-0130891.  
PR 28-APR-1999; 9905-0131449.  
PR 30-APR-1999; 9905-0132048.  
PR 30-APR-1999; 9905-0132407.  
PR 04-MAY-1999; 9905-0132484.  
PR 05-MAY-1999; 9905-0132485.  
PR 06-MAY-1999; 9905-0132486.  
PR 07-MAY-1999; 9905-0132487.  
PR 11-MAY-1999; 9905-0132863.  
PR 14-MAY-1999; 9905-0134256.  
PR 14-MAY-1999; 9905-0134219.  
PR 14-MAY-1999; 9905-0134221.  
PR 14-MAY-1999; 9905-0134370.  
PR 18-MAY-1999; 9905-0134768.  
PR 19-MAY-1999; 9905-0134941.  
PR 20-MAY-1999; 9905-0135124.  
PR 21-MAY-1999; 9905-0135353.  
PR 24-MAY-1999; 9905-0135629.  
PR 25-MAY-1999; 9905-0136021.  
PR 27-MAY-1999; 9905-0136392.  
PR 28-MAY-1999; 9905-0136782.  
PR 01-JUN-1999; 9905-0137222.  
PR 03-JUN-1999; 9905-0137528.  
PR 04-JUN-1999; 9905-0137502.  
PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139452.  
PR 16-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139492.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.



OS Homo sapiens.  
XX  
PR Key Location/Qualifiers  
XX Misc-difference 6  
FT /label= Trp or Arg  
FT  
XX  
PN W09961072-A2.  
XX  
PD 02-DEC-1999.  
XX  
PE 28-MAY-1999; 99WO-US11843.  
XX  
PR 29-MAY-1998; 98US-0087218.  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
PI Reed GL;  
XX  
DR WPI; 2000-116329/10.  
XX  
PT Inhibiting clot formation and promoting therapeutic thrombolysis using  
PT alpha-2 antiplasmin polypeptides -  
XX  
PS Claim 38; Page 32; 37pp; English.  
XX  
CC The invention relates to a novel method of detecting blood clot formation  
CC in a mammal that comprises administering a diagnostically effective  
CC amount of a detectably labeled alpha-2 antiplasmin (alpha-2AP)  
CC polypeptide. Alpha-2AP polypeptides inhibit factor XIIIa-mediated  
CC crosslinking of endogenous alpha-2AP with fibrin, to cause thrombi to  
CC undergo spontaneous physiologic lysis or accelerated lysis when  
CC administered with therapeutic thrombolytic agents. The alpha-2AP  
CC polypeptides can be labeled, and used to non-invasively detect the  
CC formation of clots and for in vivo imaging. The methods may also be used  
CC to prevent the development of blood clots in patients at risk for stroke,  
CC thrombosis and to treat patient with thrombotic condition such as stroke,  
CC myocardial infarction, pulmonary embolism, and deep venous thrombosis.  
CC The peptide may also be linked to thrombolytic agents and used to  
CC dissolve blood clots. The present sequence represents a peptide mimetic  
CC of human alpha-2AP polypeptide.  
XX  
XX  
SQ Sequence 73 AA;  
  
Query Match 1.8%; Score 7; DB 21; Length 73;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 243 GGQTALK 249  
DB 31 GGQTALK 37  
  
RESULT 40  
AAVS9187  
ID AAVS9187 standard; peptide; 73 AA.  
XX  
AC AAVS9187;  
XX  
DT 28-MAR-2000 (first entry)  
XX  
DE Alpha-2 antiplasmin (alpha-2AP) peptide mimetic.  
XX  
KW Blood clot; alpha-2 antiplasmin; alpha-2AP; factor XIIIa; fibrin;  
KW lysis; thrombolytic; thrombosis; stroke; myocardial infarction;  
KW pulmonary embolism; deep venous thrombosis.  
XX  
OS Homo sapiens.  
XX  
PN W09961072-A2.  
XX  
PD 02-DEC-1999.  
XX  
PR 28-MAY-1999; 99WO-US11843.

XX  
PR 29-MAY-1998; 98US-0087218.  
XX  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
XX  
PI Reed GL;  
XX  
DR WPI; 2000-116329/10.  
XX  
PT Inhibiting clot formation and promoting therapeutic thrombolysis using  
PT alpha-2 antiplasmin polypeptides -  
XX  
PS Claim 9; Page 29; 37pp; English.  
XX  
CC The invention relates to a novel method of detecting blood clot formation  
CC in a mammal that comprises administering a diagnostically effective  
CC amount of a detectably labeled alpha-2 antiplasmin (alpha-2AP)  
CC polypeptide. Alpha-2AP polypeptides inhibit factor XIIIa-mediated  
CC crosslinking of endogenous alpha-2AP with fibrin, to cause thrombi to  
CC undergo spontaneous physiologic lysis or accelerated lysis when  
CC administered with therapeutic thrombolytic agents. The alpha-2AP  
CC polypeptides can be labeled, and used to non-invasively detect the  
CC formation of clots and for in vivo imaging. The methods may also be used  
CC to prevent the development of blood clots in patients at risk for  
CC thrombosis and to treat patient with thrombotic condition such as stroke,  
CC myocardial infarction, pulmonary embolism, and deep venous thrombosis.  
CC The peptide may also be linked to thrombolytic agents and used to  
CC dissolve blood clots. The present sequence represents a peptide mimetic  
CC of human alpha-2AP polypeptide.  
XX  
XX  
SQ Sequence 73 AA;  
  
Query Match 1.8%; Score 7; DB 21; Length 73;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 243 GGQTALK 249  
DB 31 GGQTALK 37  
  
RESULT 41  
AAVS9188  
ID AAVS9188 standard; peptide; 73 AA.  
XX  
AC AAVS9188;  
XX  
DT 28-MAR-2000 (first entry)  
XX  
DE Alpha-2 antiplasmin (alpha-2AP) peptide mimetic.  
XX  
KW Blood clot; alpha-2 antiplasmin; alpha-2AP; factor XIIIa; fibrin;  
KW lysis; thrombolytic; thrombosis; stroke; myocardial infarction;  
KW pulmonary embolism; deep venous thrombosis.  
XX  
OS Homo sapiens.  
XX  
PN W09961072-A2.  
XX  
PD 02-DEC-1999.  
XX  
PE 28-MAY-1999; 99WO-US11843.  
XX  
PR 29-MAY-1998; 98US-0087218.  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
PI Reed GL;  
XX  
DR WPI; 2000-116329/10.  
XX  
PT Inhibiting clot formation and promoting therapeutic thrombolysis using  
PT alpha-2 antiplasmin polypeptides -

xx Claim 10; Page 29; 37pp; English.  
xx  
xx  
xx The invention relates to a novel method of detecting blood clot formation  
CC in a mammal that comprises administering a diagnostically effective  
CC amount of a detectably labeled alpha-2 antiplasmin (alpha-2AP)  
CC polypeptide. Alpha-2AP polypeptides inhibit factor XIIIa-mediated  
CC crosslinking of endogenous alpha-2AP with fibrin, to cause thrombi to  
CC undergo spontaneous physiologic lysis or accelerated lysis when  
CC administered with therapeutic thrombolytic agents. The alpha-2AP  
CC polypeptides can be labeled, and used to non-invasively detect the  
CC formation of clots and for in vivo imaging. The methods may also be used  
CC to prevent the development of blood clots in patients at risk for  
CC thrombosis and to treat patient with thrombotic condition such as stroke,  
CC myocardial infarction, pulmonary embolism, and deep venous thrombosis.  
CC The peptide may also be linked to thrombolytic agents and used to  
CC dissolve blood clots. The present sequence represents a peptide mimetic  
CC of human alpha-2AP polypeptide.  
xx  
xx  
SQ Sequence 73 AA:  
Query Match 1.8%; Score 7; DB 21; Length 73;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 243 GGQRTALK 249  
|||||  
Db 31 GGQRTALK 37  
RESULT 42  
AAU63518  
ID AUA63518 standard; Protein; 75 AA.  
xx  
xx AUA63518;  
xx  
xx 27-FEB-2002 (first entry)  
xx  
xx Propionibacterium acnes immunogenic protein #24414.  
xx  
xx SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
xx uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
xx inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
xx dermatological; osteopathic; neuroprotectant.  
xx  
xx Propionibacterium acnes.  
xx  
xx WO200181581-A2.  
xx  
xx 01-NOV-2001.  
xx  
xx 20-APR-2001; 2001WO-US12865.  
xx  
xx 21-APR-2000; 2000US-199047P.  
xx 02-JUN-2000; 2000US-208641P.  
xx 07-JUL-2000; 2000US-216747P.  
xx  
xx (CORI-) CORIXA CORP.  
xx  
xx Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A,  
xx L'maisonneuve J, Zhang Y, Jen S, Carter D.  
xx  
xx WPI; 2001-616774/71.  
xx N-PSDB; AAS59633.  
xx  
xx Propionibacterium acnes polypeptides and nucleic acids useful for  
xx treating acne vulgaris -  
xx  
xx Example 1; SEQ ID No 24713; 1069pp; English.  
xx  
xx Sequences AAU9105-AAU6017 represent propionibacterium acnes immunogenic  
xx polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
xx  
xx  
SQ Sequence 75 AA:  
Query Match 1.8%; Score 7; DB 22; Length 75;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 130 RTVKYTR 136  
|||||  
Db 27 RTVKYTR 33  
RESULT 43  
AAM84718  
ID AAM84718 standard; Protein; 86 AA.  
xx  
xx AAM84718;  
xx  
xx 07-NOV-2001 (first entry)  
xx  
xx Human immune/haematopoietic antigen SEQ ID NO:12311.  
xx  
xx Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
xx cytostatic; gene therapy; vaccine; metastasis.  
xx  
xx Homo sapiens.  
xx  
xx WO200157182-A2.  
xx  
xx 09-AUG-2001.  
xx  
xx 17-JAN-2001; 2001WO-US01354.  
xx  
xx 31-JAN-2000; 2000US-0179065.  
xx 04-FEB-2000; 2000US-0180628.  
xx 24-FEB-2000; 2000US-0184664.  
xx 02-MAR-2000; 2000US-0186350.  
xx 16-MAR-2000; 2000US-0189874.  
xx 17-MAR-2000; 2000US-0190076.  
xx 18-APR-2000; 2000US-0198123.  
xx 19-MAY-2000; 2000US-0205515.  
xx 07-JUN-2000; 2000US-0209467.  
xx 28-JUN-2000; 2000US-0214886.  
xx 30-JUN-2000; 2000US-0215135.  
xx 07-JUL-2000; 2000US-0216647.  
xx 07-JUL-2000; 2000US-0216880.  
xx 11-JUL-2000; 2000US-0217487.  
xx 11-JUL-2000; 2000US-0217496.  
xx 14-JUL-2000; 2000US-0218290.  
xx 26-JUL-2000; 2000US-0220963.  
xx 26-JUL-2000; 2000US-0220964.  
xx 14-AUG-2000; 2000US-0224518.  
xx 14-AUG-2000; 2000US-0224519.  
xx 14-AUG-2000; 2000US-0225213.  
xx 14-AUG-2000; 2000US-0225214.





SQ Sequence 86 AA: 1.8%; Score 7; DB 22; Length 86;  
Query Match Best Local Similarity 100.0%; Pred. No. 87;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 43 LIVGRT 49  
|||||  
DB 6 LIVGRT 12

RESULT 44  
ABG17892  
ID ABG17892 standard; Protein; 101 AA.  
XX  
AC ABG17892;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #17883.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW Food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
XX 23-AUG-2000; 2000US-0649167.  
PA (HYSE-) HYSEQ INC.  
PI Drmanac RT, Liu C, Tang YT;  
XX N-PSDB; AAS82079.  
DR WPI: 2001-639362/73.  
DR N-PSDB; AAS82079.

New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
PS Claim 20: SEQ ID NO 48251; 103pp; English.

The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 101 AA:

Query Match 1.8%; Score 7; DB 22; Length 101;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 173 VFAVASE 179  
|||||  
DB 83 VFAVASE 89

RESULT 45  
AAG05445  
ID AAG05445 standard; Protein; 106 AA.  
XX  
AC AAG05445;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1854.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135533.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151067.  
PR 30-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0161405.  
PR 22-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 7; DB 21; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KDAIEKL 127  
DB 98 KDAIEKL 104

Search completed: November 5, 2002, 16:22:40  
Job time : 67 secs

---

